Multidrug resistance and minimal residual disease in childhood malignancies

This thesis is submitted as fulfillment of the requirements for the degree of Ph.D. in Medical Sciences by ir. Katrien Swerts, 2005

Promotors
Prof. Dr. Jan Philippé
Prof. Dr. Geneviève Laureys

Department of Pediatric Hematology and Oncology
Ghent University Hospital, De Pintelaan 185, 9000 Ghent, Belgium
++32/9/240.66.42 (Phone)
++32/9/240.49.85 (Fax)
Katrien.Swerts@UGent.be
Thesis submitted to fulfill the requirements for the degree of Doctor in Medical Sciences

June 2005

Promotors:

Prof. Dr. Jan Philippé
Ghent University, Belgium

Prof. Dr. Geneviève Laureys
Ghent University, Belgium

Members of the examination committee:

Prof. Dr. Jean Bénard
Institut Gustave-Roussy, France

Prof. Dr. Andy Pearson
University of Newcastle upon Tyne, UK

Prof. Dr. Leroux-Roels
Ghent University, Belgium

Prof. Dr. Marc Bracke
Ghent University, Belgium

Prof. Dr. Frank Speleman
Ghent University, Belgium

Prof. Dr. Yves Benoit
Ghent University, Belgium

Prof. Dr. Lucien Noens
Ghent University, Belgium

The research described in this thesis was conducted in the Department of Clinical Chemistry, Microbiology and Immunology, Ghent University Hospital, Ghent, Belgium

Research funded by a Ph.D. grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen) and the Children’s Cancer Fund 'Kinderkankerfonds'.
# Table of contents

## INTRODUCTION

1. CHILDHOOD MALIGNANCIES
   1.1. INTRODUCTION
   1.2. ACUTE LYMPHOBLASTIC LEUKEMIA
      1.2.1. CLINICAL FEATURES
      1.2.2. CLASSIFICATION
      1.2.3. PROGNOSTIC FACTORS AND THERAPY
   1.3. NEUROBLASTOMA
      1.3.1. CLINICAL FEATURES
      1.3.2. CLASSIFICATION
      1.3.3. HISTOPATHOLOGY
      1.3.4. GENETICS
      1.3.5. THERAPY AND PROGNOSIS

2. AIMS OF THIS THESIS

## MULTIDRUG RESISTANCE

1. INTRODUCTION

2. CLINICAL RELEVANCE OF MULTIDRUG RESISTANCE IN CHILDHOOD ALL
   2.1. PAPER 1
   2.2. PAPER 2

3. EVALUATION OF P-GLYCOPROTEIN ACTIVITY
   3.1. PAPER 3

4. DISCUSSION AND FUTURE PERSPECTIVES
MINIMAL RESIDUAL DISEASE ............................................................................................71
1. INTRODUCTION ..........................................................................................................73
2. DETECTION OF DISSEMINATED NEUROBLASTOMA CELLS .................................74
   2.1. PAPER 4 .............................................................................................................74
   2.2. PAPER 5 .............................................................................................................92
   2.3. PAPER 6 ............................................................................................................104
   2.4. PAPER 7 ............................................................................................................126
3. CLINICAL IMPORTANCE OF NEUROBLASTOMA MICROMETASTASES ..........151
   3.1. STAGING AND RISK ASSESSMENT AT DIAGNOSIS ........................................151
   3.2. MONITORING DISEASE COURSE AND RESPONSE TO THERAPY ....................151
   3.3. SCREENING OF AUTOLOGOUS STEM CELL PREPARATIONS ..........................152
4. DISCUSSION AND FUTURE PERSPECTIVES ........................................................153

REFERENCES ......................................................................................................................157

SUMMARY ..........................................................................................................................173

SAMENVATTING .................................................................................................................175

RESUME .............................................................................................................................177

ABBREVIATIONS ...............................................................................................................179

DANKWOORD .....................................................................................................................181

CURRICULUM VITAE .........................................................................................................183
Chapter 1

Introduction
1. Childhood malignancies

1.1. Introduction

During the past decades, remarkable advances have been achieved in the treatment and survival of children with cancer. Despite these important advances, cancer remains the second leading cause of death in children, after mortality related to accidents. The types, distribution, prognosis and responsiveness to therapy of the malignancies occurring in the pediatric age group differ markedly from those that occur in adult populations. Primary epithelial tumors, originating from cells covering internal and external body surfaces, predominate in adults (e.g. cancer of the colon, lung and breast). By contrast, childhood tumors are usually of mesenchymal origin or derive from the hematopoietic or central nervous systems. Acute leukemia is the most common cancer in children younger than 15 years of age, followed by tumors of the central nervous system (CNS), lymphoma and neuroblastoma (Figure 1).

Figure 1: Distribution of cancer in children younger that 15 years of age in Belgium

(Based on data from the National Cancer Registry, 1998)
1.2. Acute Lymphoblastic Leukemia

1.2.1. Clinical features

Acute lymphoblastic leukemia (ALL) is the most common malignancy in children. It accounts for one fourth of all childhood cancers and approximately 75% of all childhood leukemias. The clinical incidence of ALL in the United States is estimated at 3 to 4 per 100,000 Caucasian children younger than 15 years of age. ALL is more common among whites than blacks, but a racial predilection has not been demonstrated. Boys have higher incidence rates than girls with a male-to-female ratio of 1.3. Childhood ALL has a peak incidence between 2 and 6 years in white children.

Various constitutional chromosomal abnormalities are associated with childhood leukemia. Children with Down’s syndrome have a 15-fold greater risk of developing leukemia than normal children. In addition, patients with fragile chromosomes, as in Bloom’s syndrome, Fanconi’s anemia and ataxia telangiectasia, are at high risk to develop ALL. The frequency of leukemia is higher than expected in families of leukemia patients and the concordance of acute leukemia in monozygotic twins under the age of 7 is estimated at 25%. Both in utero exposure to diagnostic X-rays and postnatal exposure to therapeutic doses of ionizing radiation can mediate the development of ALL. However, the actual percentage of leukemias directly attributable to radiation is probably small. So far, no association between chronic exposure to electromagnetic fields or chemicals and ALL has been found. Some reports have suggested that prenatal or postnatal exposure to viruses is associated with a higher risk of ALL. However, no definitive link between viral infections and the occurrence of ALL has been established. Congenital immunodeficiency diseases and abnormalities of the immune system are occasionally observed in pediatric ALL patients.

The molecular basis for leukemic transformation in humans is unknown. It has been suggested that acute leukemia develops as a result of spontaneous mutations in a single hematopoietic progenitor cell leading to dysregulated growth and arrested differentiation. Proposed molecular mechanisms for leukemic induction include activation of oncogenes or creation of fusion genes with oncogenic properties and/or loss or inactivation of tumor suppressor genes.

A complete morphologic, immunologic and genetic examination of bone marrow aspirates is necessary to establish the diagnosis of ALL. Most centers require 20% to 30% leukemic cells in the bone marrow before the diagnosis is confirmed.
Initial blood counts show a broad spectrum of abnormal findings. Approximately 20% of children present with leukocyte counts greater than $50 \times 10^3/\mu l$. Neutropenia (< 500 granulocytes/µl) and thrombocytopenia (< $100 \times 10^3$ platelets/µl) are common phenomena and approximately 80% of the patients present with anemia (hemoglobin < 10g/dl). A variety of other abnormal laboratory findings may be encountered. Increased serum uric acid levels reflect increased anabolism and catabolism of purines, hypercalcemia may result from leukemic infiltration in the bone, and elevated serum lactate dehydrogenase (LDH) and phosphorus levels are associated with tumor cell lysis. Approximately 5% to 10% of de novo ALL patients present with an anterior mediastinal mass which can be detected on chest radiographs. Bone radiography may disclose bone involvement. Cerebrospinal fluid should be examined because it contains leukemic blasts in 3% to 5% of leukemic patients.

### 1.2.2. Classification

Formerly, ALL was classified into three categories (L1, L2 and L3) according to the system proposed by the French-American-British (FAB) cooperative group.\(^{19}\) This FAB system was based solely on morphological and cytochemical findings. The discovery of new genetic aberrations, predicting clinical behavior and outcome, necessitated the development of a new classification scheme. The World Health Organization (WHO) defined a new system based on genetic, cytochemical and morphological information.\(^{20}\) ALLs are divided into two major categories: precursor B-cell lymphoblastic leukemia (B-ALL) and precursor T-cell lymphoblastic leukemia (T-ALL).

#### 1.2.2.1. Morphology

Precursor B-ALL lymphoblasts in bone marrow smears or imprints vary from small cells with scant cytoplasm, condensed nuclear chromatin and indistinct nucleoli to larger cells with moderate amounts of cytoplasm, dispersed nuclear chromatin and multiple, variably prominent nucleoli.\(^{20}\) In 10% of the cases, azurophilic granules are present.

T-ALL cells are cytomorphologically quite similar to precursor B lymphoblasts. In smears, precursor T-ALL cells are of medium size with high nuclear/cytoplasmic ratio. There may be a considerable difference in size, ranging from small lymphoblasts with very condensed nuclear chromatin and indistinct nucleoli to larger blasts with finely dispersed chromatin and relatively prominent nucleoli.\(^{20}\)
1.2.2.2. **Cytochemistry**

Lymphoblasts are negative for myeloperoxidase and Sudan Black-B, but may react with periodic acid Schiff (PAS) or non-specific esterase.\(^\text{20}\) T-ALL cells often show focal acid phosphatase activity in bone marrow smears or imprints.

1.2.2.3. **Immunophenotype**

Precursor B-ALL cells express terminal deoxynucleotidyl transferase (TdT) and HLA-DR antigens.\(^\text{20}\) They are almost always positive for CD19, cytoplasmic (c) CD79a, CD10 and CD22. CD20 is expressed in a minority of cases and CD45 may be absent. The myeloid associated antigens CD13 and CD33 are frequently expressed as aberrant markers. Based on their immunophenotype, precursor B-ALL blasts are divided into three subgroups: early precursor B-ALL, common B-ALL and pre-B-ALL (Figure 2). Early precursor B-ALL blasts express CD34, CD19, cCD79a, CD22 and nuclear TdT. The more mature common B-ALL cells are CD10 positive and the most differentiated pre-B-ALL cells express cytoplasmic μ chains. Surface immunoglobulin is characteristically absent.

Precursor T-ALL cells are TdT positive and variably express CD1a, CD2, CD3, CD4, CD5, CD7 and CD8.\(^\text{20}\) Of these, CD7 and cCD3 are most frequently positive. CD4 and CD8 are often co-expressed on the blasts and CD10 may be positive. In addition to CD117, myeloid markers CD13 and/or CD33 are occasionally present. T lymphoblasts may present clonal rearrangements of the T-cell receptor genes. Immunoglobulin gene rearrangements are also observed in T-ALL cells. Precursor T-ALL is classified into three subtypes which represent different stages of intrathymic differentiation (Figure 2): immature T-ALL, common T-ALL and mature T-ALL.

1.2.2.4. **Genetics**

At the moment, genetic abnormalities can be identified in approximately 65% to 70% of pediatric ALL patients.\(^\text{20,21}\) They are prognostically important and are used to modify treatment. Patients with hyperdiploid chromosome numbers (between 51 and 65 chromosomes or DNA-index between 1.16 and 1.6) have a good prognosis.\(^\text{22-25}\) In addition, \(t(12;21)(p13;q22)\) is also associated with favorable outcome.\(^\text{26-28}\) This rearrangement results in the fusion of the *TEL* gene (12p13), a member of the ETS-like family of transcription factors, and the transcription factor-encoding *AML1* gene (21q22). The \(t(12;21)\) is detected in about 20% to 30% of children with precursor B-ALL.\(^\text{29,30}\)
Other cytogenetic findings, including t(9;22)⁵¹-⁵⁴, t(4;11)⁵⁵-⁵⁹, t(1;19)⁶⁰-⁶², and hypodiploidy⁶³,⁶⁴, are associated with poor prognosis. The t(9;22) transposes the 3’ portion of the cytoplasmic tyrosine kinase gene ABL (9q34) to the 5’ region of BCR (22q11.2), generating a chimeric tyrosine kinase oncogene.⁶⁵ The translocation is present in 3% to 5% of pediatric ALL patients.

Cytogenetic abnormalities involving the MLL gene on 11q23 are frequently observed in childhood ALL. The most common is t(4;11)(q21;q23). This translocation is usually found in early precursor B-ALL patients. Translocation t(1;19), found in 25% of children with pre-B-ALL, results in the fusion of the transcription activator-encoding gene E2A (19p13.3) with the homeobox gene PBX1 (1q23).⁶⁶,⁶⁷ Genetic abnormalities including del(6q)⁶⁸,⁶⁹, del(9p)⁷⁰, del(12p)⁷¹,⁷², and hyperdiploidy (less than 51 chromosomes)⁷³-⁷⁵ are associated with an intermediate prognosis.

Figure 2: Immunophenotype of acute lymphoblastic leukemia
Introduction

In contrast to precursor B-ALL, the recurrent chromosomal alterations observed in precursor T-ALL have not been helpful in providing significant prognostic information.\textsuperscript{56,57} In about one third of the T-ALL patients translocations involving the $\alpha$ and $\delta$ T-cell receptor loci at 14q11.2, the $\beta$ locus at 7q35 or the $\gamma$ locus at 7p14-15 have been detected.\textsuperscript{58,59} Partner genes include the transcription factors $\text{MYC (8q24.1), TAL1 (1p32), RBTN1 (11p15), RBTN2 (11p13)},$ and $\text{HOX11 (10q24)}$ and the cytoplasmic tyrosine kinase $\text{LCK (1p34.3-35).}$\textsuperscript{60,61} In 5% of the T-ALL patients, the $\text{TAL1}$ gene is dysregulated by a deletion in its 5' regulatory region.\textsuperscript{62} The deletion del(9p), resulting in loss of the tumor suppressor gene CDKN2A (INK4A/p16), is frequently observed in T-ALL patients.\textsuperscript{63}

1.2.3. Prognostic factors and therapy

The ability to identify subsets of patients who have a better or worse prognosis than average has contributed to the success of modern ALL treatment programs. Various clinically significant features, including initial leukocyte count, age at diagnosis, sex, cytogenetic aberrations, immunophenotype, and response to therapy, have been identified.\textsuperscript{1,2,64} (Table 1)

It is important to realize that the impact of prognostic factors depends on the efficacy of the therapy. Some parameters have lost their clinical relevance (e.g. race) due to the major improvements in the treatment of ALL patients.

Children are divided into different risk groups based on these prognostic parameters. Low risk patients are treated with less intensive and toxic chemotherapy, whereas high risk patients receive a more aggressive treatment. So far, no single classification system is universally accepted but the initial leukocyte count and age at diagnosis form the basis for most methods of prognostic stratification.

Although specific approaches to therapy might differ from center to center, all modern treatment regimens comprise four major elements: remission induction, intensification/consolidation, CNS-directed therapy and continuation of therapy.\textsuperscript{1,2,65-67} Induction therapy typically includes administration of a glucocorticoid (dexamethasone or prednisone), vincristine, and L-asparaginase. The leukemic blasts of ALL patients are usually sensitive to chemotherapy at the time of diagnosis. Consequently, complete remission (defined as the absence of clinical signs and symptoms, the recovery of normal blood cell counts, and the recovery of normocellular bone marrow with less than 5% blast cells) will be achieved in 97% to 98% of children.
Table 1  
Factors associated with prognosis for children with ALL

<table>
<thead>
<tr>
<th></th>
<th>Favorable Prognosis</th>
<th>Unfavorable Prognosis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial Leukocyte Count</strong></td>
<td>≤ 50.10³/µl</td>
<td>&gt; 50.10³/µl</td>
</tr>
<tr>
<td><strong>Age at Diagnosis</strong></td>
<td>2-10 years</td>
<td>&lt; 2 years (especially infants) &gt; 10 years</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td><strong>Cytogenetic Aberrations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>✓ <strong>Ploidy</strong></td>
<td>Hyperploidy</td>
<td>Hypoploidy Pseudodiploidy Tetraploidy</td>
</tr>
<tr>
<td>✓ <strong>Chromosomal rearrangements</strong></td>
<td>t(12;21)</td>
<td>t(9;22), t(4;11), t(1;19)</td>
</tr>
<tr>
<td><strong>Immunophenotype</strong></td>
<td>Early Precursor B-ALL CD10⁺</td>
<td>T-ALL (especially CD3⁺) Pre-B-ALL</td>
</tr>
<tr>
<td><strong>Response to Treatment/MRD</strong></td>
<td>Early response (&lt; 1000 blast in peripheral blood after one week of prednisolone treatment. Cytomorphologic remission at day 35)</td>
<td>Late response</td>
</tr>
</tbody>
</table>

The intensification or consolidation therapy aims at the destruction of residual lymphoblasts to a level comparable with cure. During this consolidation period, patients receive either increased doses of agents used previously or entirely new agents with minimal cross-resistance. The importance of preventive CNS therapy was first demonstrated by investigators at St. Jude Children’s Research Hospital. CNS-directed therapy is essential because leukemia in the CNS reduces the chance of cure and increases morbidity related to the need for intensive remission-retrieval therapy. Most children with ALL are treated with methotrexate or triple-agent chemotherapy (methotrexate, corticoids, and cytarabine). ALL patients require prolonged continuation treatment. The most common continuation therapy relies on the daily administration of 6-mercaptopurine and weekly administration of methotrexate. The general rule in modern protocols is to discontinue all therapy in patients in whom continuous complete remission has been achieved for 2 years.
Introduction

Long-term event-free survival rates for newly diagnosed children with ALL have improved markedly during the past 30 years. However, 20% to 30% of the pediatric ALL patients relapse and cannot be cured with modern therapy. Intrinsic or acquired resistance to a wide variety of structurally and functionally unrelated chemotherapeutic compounds is one of the most important causes of treatment failure in childhood ALL. A variety of cellular mechanisms can give rise to multidrug resistance (MDR), but best studied is the transmembrane protein-mediated efflux of cytotoxic compounds which leads to a decreased cellular drug accumulation and toxicity. The clinical significance of multidrug resistance-related proteins in childhood ALL and the laboratory techniques used to study these efflux pumps are discussed in more detail in Chapter 2.
1.3. Neuroblastoma

1.3.1. Clinical features

Neuroblastoma is the most common extracranial solid tumor in children and accounts for 8% to 10% of all childhood malignancies.\textsuperscript{100} The clinical incidence of neuroblastoma in the Western world is estimated at 1 per 100,000 children younger than 15 years of age.\textsuperscript{101,102} The median age at diagnosis is 22 months and 97% of all neuroblastoma patients are less than 10 years old. The tumor is slightly more common in boys than in girls with a male-to-female ratio of 1.3/1.\textsuperscript{103}

The etiology of neuroblastoma is unknown. So far, no clear association between the incidence of neuroblastoma and prenatal or postnatal exposure to drugs, chemicals, electromagnetic fields or radiation has been found.\textsuperscript{104-106} A few families with more than one affected sibling have been reported, but a genetic basis has not yet been demonstrated.\textsuperscript{107,108}

Neuroblastomas are neuroectodermal tumors which arise from primitive, pluripotent, sympathetic cells or sympathogonia. These embryonic cells originate from the neural crest and form, after migration and differentiation, the different normal tissues of the sympathetic nervous system (Figure 3).\textsuperscript{109} As a consequence, most neuroblastomas occur within the adrenal gland (35%) or along the sympathetic nervous system chain ganglia (abdomen (30%), thorax (20%)).\textsuperscript{100} The location of the primary tumor has some impact on outcome. Children with a primary tumor in the thorax or pelvis have a better prognosis than those with adrenal tumors.\textsuperscript{110}

The diagnosis of neuroblastoma is established by pathologic examination of tumor tissue by light microscopy.\textsuperscript{111} Alternatively, the diagnosis is made based on the detection of neuroblastoma cells in bone marrow aspirates or trephine biopsies in combination with the demonstration of increased urinary catecholamine metabolites (e.g. homovanillic acid (HVA) and vanillylmandelic acid (VMA)). Plain radiography, bone scintigraphy, ultrasound, computed tomography (CT-scan), magnetic resonance imaging (MRI-scan) and meta-iodobenzylguanidine (MIBG) scintigraphy are used to determine the extend of the disease.
The evaluation of bone marrow smears is not only important for clinical staging and risk assessment at diagnosis, it can also be used to monitor therapeutic response during treatment. Conventional cytology of bone marrow smears is still the only accepted technique for the detection of disseminated neuroblastoma cells. However, the sensitivity of this method is limited since a tumor cell number below 0.1 % cannot be detected by conventional cytomorphology. Therefore, more sensitive and specific detection techniques are urgently needed. New laboratory methods used to screen for residual neuroblastoma cells, will be discussed in Chapter 3.

1.3.2. Classification

In the past, three major staging systems were used to classify neuroblastoma patients in different risk groups: (1) system proposed by Evans (Children’s Cancer Study Group); system used by the Pediatric Oncology Group; and (3) system utilized by the American Joint Committee on Cancer. The differences between these staging systems were substantial and made the comparison of results from clinical trials and biological studies impossible. Therefore, a new staging system, based on the Evans classification, was formulated by Brodeur and colleagues. This International Neuroblastoma Staging System (INSS) is generally accepted and classifies patients based on clinical, radiographic and surgical evaluation (Table 2).
The most important prognostic parameters are disease stage and age of the patient at diagnosis.\textsuperscript{119} Children over the age of one year have a worse prognosis than infants (< 1 year) and there is a progressive decline in outcome with increasing stage of disease. The exceptions to this are infants with stage 4S disease. Their outcome is similar to that of patients with stage 1 or 2 disease.

\subsection*{1.3.3. Histopathology}

Peripheral neuroblastic tumors belong to the group of "small blue round cell" neoplasms of childhood. Other members of this group are lymphoblastic lymphomas, Ewing’s sarcomas, primitive neuroectodermal tumors, rhabdomyosarcomas and other undifferentiated small cell soft tissue sarcomas.\textsuperscript{100} Historically, peripheral neuroblastic tumors have been divided into three basic morphologic categories, based on the presence of neuroblasts and/or gangliocytic differentiated cells: neuroblastoma, ganglioneuroblastoma, and ganglioneuroma.\textsuperscript{100} Recently, a new prognostically significant and biologically relevant classification system has been proposed by the International Neuroblastoma Pathology Committee (INPC).\textsuperscript{120} This INPC classification system is exclusively based on morphologic changes and focuses not only on the gangliocytic differentiation of neuroblasts, but also on the relative amount of Schwannian cell stroma, as originally proposed by Shimada.\textsuperscript{121} Peripheral neuroblastic tumors are divided into four basic

\begin{table}[h]
\centering
\begin{tabular}{|l|p{13cm}|}
\hline
\textbf{Stage} & \textbf{Description} \\
\hline
1 & Localized tumor confined to the area of origin; complete gross excision, with or without microscopic residual disease; identifiable ipsilateral and contralateral lymph nodes negative microscopically \\
2A & Unilateral tumor with incomplete gross excision; identifiable ipsilateral and contralateral lymph nodes negative microscopically \\
2B & Unilateral tumor with complete or incomplete gross excision; positive ipsilateral regional lymph nodes; identifiable contralateral lymph nodes negative microscopically \\
3 & Tumor infiltrating across the midline with or without regional lymph node involvement; or unilateral tumor with contralateral regional lymph node involvement; or midline tumor with bilateral lymph node involvement \\
4 & Dissemination of tumor to distant lymph nodes, bone, bone marrow, liver, or other organs (except as defined in stage 4S) \\
4S & Localized primary tumor as defined for stage 1 or 2 with dissemination limited to liver, skin, or bone marrow \\
\hline
\end{tabular}
\caption{International Neuroblastoma Staging System\textsuperscript{111,118}}
\end{table}
morphological categories and their subtypes: 1) neuroblastoma (Schwannian stroma-poor), undifferentiated, poorly differentiated, and differentiating; 2) ganglioneuroblastoma, intermixed (Schwannian stroma-rich); 3) ganglioneuroma (Schwannian stroma-dominant); and 4) ganglioneuroblastoma, nodular (composite Schwannian stroma-rich/stroma-dominant and stroma-poor).

Neuroblastomas are composed of uniform, small, round cells with scant cytoplasm and an often granular (“salt and pepper”) chromatin structure. Depending on the presence of neuropil and gangliocytic differentiated tumor cells, undifferentiated, poorly differentiated and differentiating subtypes are recognized. Scattered Schwann cells can be detected in delicate fibrovascular septa. The tumor cells may form Homer Wright rosettes (clusters of neuroblasts around a core of neuropil). The fully differentiated and benign counterpart of neuroblastoma is ganglioneuroma. It is composed of mature ganglion cells embedded in bulky stroma composed of Schwann cells. The differentiation of ganglioneuroblastomas is intermediate to neuroblastomas and ganglioneuromas. They present with large areas of Schwannian cell stroma which surrounds islands of stroma-poor neuroblastic tissue, either as small scattered foci (intermixed type) or as grossly visible nodules (nodular type).

Based on patient age and morphological features (e.g. grade of neuroblastic differentiation, presence or absence of Schwannian stroma and mitosis-karyorrhexis index (MKI)) peripheral neuroblastic tumors are classified into favorable histology or unfavorable histology. Poorly differentiated and differentiating neuroblastomas with low and intermediate MKI below 1.5 years of age, and differentiating neuroblastoma with low MKI between 1.5 and 5 years of age are categorized as favorable. All other neuroblastomas are considered unfavorable. Nodular ganglioneuroblastomas are categorized according to the histology of the neuroblastic nodule based on the same criteria as for neuroblastomas. Intermixed ganglioneuroblastomas and ganglioneuroma are considered favorable.

1.3.4. Genetics

1.3.4.1. Ploidy

The DNA content of neuroblastoma cells can be used to predict response to treatment and outcome of patients, especially in infants less than 1 year of age with advanced stage disease. Based on their ploidy, neuroblastomas can be divided into two distinct subgroups. The majority of the neuroblastoma cell lines, xenografts and advanced stage tumors have near-diploid or near-tetraploid chromosome numbers, whereas favorable neuroblastomas, especially those from infants, are characterized by a hyperdiploid or near-triploid DNA content. Virtually all near-diploid or near-tetraploid tumors show various genomic
Introduction

aberrations, including amplifications, deletions, and unbalanced translocations as a result of genomic instability. This is not the case in hyperdiploid or near-triploid tumors. These tumors have a basic defect in the machinery of mitosis and chromosome segregation and are characterized by whole chromosome gains with few if any structural rearrangements. Look and coworkers were the first to describe a relationship between hyperdiploidy and better response to chemotherapy in infants with unresectable neuroblastoma. In a subsequent Pediatric Oncology Group study, hyperdiploidy was clearly associated with long term survival in infants whereas diploidy predicted early treatment failure. However, there was no relationship between ploidy and clinical outcome in children over two years of age.

In general, patients with near-triploid tumors usually have favorable clinical and biological prognostic factors and excellent survival rates, whereas unfavorable prognostic factors and a poor outcome were observed in patients who had near-diploid or near-tetraploid tumors.

1.3.4.2. MYCN amplification

The amplification of MYCN, a proto-oncogene located on the distal short arm of chromosome 2 (2p24.1), is one of the most intriguing and striking genetic aberrations found in neuroblastoma. The extra MYCN copies are located on extrachromosomal double minutes (dmins) or homogeneously staining regions (HSRs) involving different chromosomes. Amplification values in neuroblastoma may range from 10 to more than 500-fold but values around 50 to 100-fold are generally seen in tumors. Brodeur and coworkers were the first to study the frequency and significance of MYCN amplification in neuroblastoma. MYCN amplification was found in 25% of primary tumors from untreated patients and was strongly associated with advanced disease, rapid tumor progression and poor outcome. MYCN amplification is virtually always associated with very high MYCN expression levels. However, the independent prognostic significance of high MYCN mRNA or protein levels remains unclear.

The amplified DNA encompassing MYCN is quite large (500 to 1 000 kb). Therefore, it has been suggested that additional genes, located near MYCN, might be co-amplified. However, so far no other, consistently co-amplified gene has been identified. In approximately 40% to 50% of MYCN-positive neuroblastomas, the RNA helicase gene DDX1 is co-amplified. Similarly, a gene termed NAG (neuroblastoma amplified gene) can be co-amplified with MYCN. However, the extent to which these co-amplified genes contribute to the neuroblastoma phenotype remains a matter for debate.
1.3.4.3. Deletion of the short arm of chromosome 1

Deletion of the short arm of chromosome 1 is a common cytogenetic abnormality in neuroblastoma. Loss of heterozygosity (LOH) is found in 30% to 50% of primary tumors. Several regions of deletion have been identified. A common deleted region is located at the distal end of chromosome 1p (1p36.2 to 1p36.3). Schleiermacher and colleagues reported three cases with interstitial deletions of 1p32 only and Takeda and coworkers divided their tumors with 1p deletion into two groups: those with large terminal deletions which encompassed the region from 1pter to 1p32 and those with smaller interstitial deletions in 1p36. They found a relationship between large terminal 1p deletions, MYCN amplification and unfavorable outcome. Caron and colleagues confirmed these results. These findings suggest that at least two neuroblastoma suppressor genes are located on 1p (one at 1p36.3 and another one at 1p35-p36.1).

Although there is a strong correlation between 1p deletion and high-risk features such as age above 1 year at diagnosis, metastatic disease and MYCN amplification, the independent prognostic significance of 1p LOH remains unclear. A retrospective Children’s Cancer Group study indicated that 1p36 LOH was an independent predictor of event-free survival but not of overall survival. These findings suggest that 1p LOH analysis may be used to identify low- or intermediate-risk patients who are likely to relapse.

1.3.4.4. Gain of the long arm of chromosome 17

Unbalanced translocations resulting in gain of 17q material and partial loss of the partner chromosome are the most common genetic aberrations in primary neuroblastomas. The most frequent partner is chromosome 1p but other partner chromosomes may also be involved. Gain of 17q21-qter is found in 50% to 75% of primary tumors and is associated with a poor prognosis. The smallest region of gain is a 25 Mb region at 17q23.1-17qter. This region contains probably one or multiple genes that contribute to neuroblastoma tumorigenesis when present in increased copy number. So far, no candidate genes have been identified.

1.3.5. Therapy and Prognosis

Neuroblastomas are often described as enigmatic and unpredictable because of their heterogeneous biologic, genetic, and morphologic characteristics. Some neuroblastomas are highly malignant and respond poorly to current treatment whereas others regress spontaneously or mature to benign ganglioneuroma. Based on age, stage and MYCN status, physicians are able to predict, to some extent, the biological course of the disease.
Three different risk groups are defined. Patients with stage 1, 2 and 4S disease without MYCN amplification are included in the low-risk group (36% of all patients). They require minimal therapy and are usually treated with surgery only. Their prognosis is very good with a 5-year survival rate of approximately 96%. The group of the intermediate-risk patients includes patients with stage 3 neuroblastoma and infants with stage 4 disease without MYCN amplification (10% of all children). These patients are treated with conventional chemotherapy in combination with surgery and have a 74% survival rate. The third group includes high-risk stage 4 patients over the age of one year, regardless of their MYCN status in addition to children with MYCN-amplified stage 1, 2, 3 and 4S disease (54% of all patients). When chemotherapy, surgical resection of the primary tumor, myeloablative therapy with hematopoietic stem cell rescue and therapy with 13-cis-retinoic acid are combined, a survival rate of approximately 30% to 40% is obtained.¹⁵⁷

Recently, a new model for improved risk assessment and therapy stratification has been proposed by Vandesompele and colleagues.¹⁵⁸ Based on comparative genomic hybridization data from 231 untreated neuroblastomas, three major clinicogenetic subgroups have been identified. Subgroup 1 consists of infants with a low stage near-triploid tumor with numerical aberrations and favorable histology. These patients have an excellent outcome. The second group represents older children with a near-diploid/tetraploid stage 4 tumor, characterized by unfavorable histology, partial 17q gain, 11q loss and/or 3p deletion. These children have a poor overall 5-year survival of approximately 40%. Patients belonging to subgroup 3 have a comparable survival rate. They present with a near-diploid/tetraploid high stage tumor, displaying unfavorable histology, partial 17q gain, 1p loss and MYCN amplification. Patients belonging to subgroup 2 are generally older compared to children from subgroup 3 (median age at diagnosis of 41 months compared to 26 months).

Gene expression profiling can also be used to identify genetic signatures predictive of clinical outcome.¹⁵⁹-¹⁶³ Several research groups compared expression profiles from tumors with favorable and unfavorable characteristics in order to identify differentially expressed genes which could be new prognostic markers or therapeutic targets in aggressive neuroblastoma.¹⁵⁹-¹⁶¹ Krasnoselsky and colleagues identified a specific subset of cell cycle and/or chromosome segregation genes that distinguish stage 4 tumors from all lower stage neuroblastomas.¹⁶³ Wei et al. identified 19 genes (e.g. MYCN, CD44, DLK1, SLIT3 and ROBO2) which predicted outcome for 98% of the studied neuroblastoma patients. However, integration of these findings into clinical risk prediction algorithms requires independent testing in large groups of uniformly treated patients.
2. Aims of this thesis

Although the prognosis of children diagnosed with ALL or neuroblastoma has improved markedly during the past four decades, a considerable number of children relapses and cannot be cured with current chemotherapy. One of the major goals of researchers, working in the field of pediatric hematology-oncology, is the identification of new prognostic parameters which can be used to classify children into different risk groups. Low risk patients are treated with less intensive and toxic chemotherapy, whereas high risk patients receive a more aggressive treatment. Recently, two approaches have become popular: 1) screening for transmembrane protein-mediated multidrug resistance (MDR) in childhood ALL and 2) minimal residual disease (MRD) analysis in neuroblastoma. However, the clinical relevance of these prognostic parameters remains controversial due to the lack in sensitive and specific detection methods. Consequently, we decided to develop, optimize and standardize new detection techniques for the assessment of MDR and MRD. In addition, the clinical relevance of these parameters in childhood malignancies was studied. Briefly, the aims of the thesis are:

- Review current data on the clinical significance of multidrug resistance-related proteins in childhood ALL.
- Evaluation of the prognostic significance of P-glycoprotein (P-gp) expression and activity in childhood ALL.
- Development of a highly sensitive functional flow cytometric assay to measure P-gp activity in pediatric ALL patients.
- Development and standardization of an immunocytochemical staining method for detection of disseminated neuroblastoma cells (in collaboration with the SIOPEN Bone Marrow Subcommittee).
- Development of multiparameter flow cytometric assays to screen for occult neuroblastoma cells. Comparison of the flow cytometric results with those of an immunocytochemical reference assay.
- Development of molecular methods to detect rare tumor cells and study of the prognostic significance of MRD in neuroblastoma.
- Review current knowledge about the detection techniques used to study MRD in neuroblastoma.
Chapter 2

Multidrug Resistance
MULTIDRUG RESISTANCE .......................................................................................................................... 19
1. INTRODUCTION ............................................................................................................................................. 21
2. CLINICAL RELEVANCE OF MULTIDRUG RESISTANCE IN CHILDHOOD ALL ............................................. 22
   2.1. PAPER 1 ....................................................................................................................................................... 22
   2.2. PAPER 2 ....................................................................................................................................................... 46
3. EVALUATION OF P-GLYCOPROTEIN ACTIVITY .............................................................................................. 58
   3.1. PAPER 3 ....................................................................................................................................................... 58
4. DISCUSSION AND FUTURE PERSPECTIVES .................................................................................................. 67
1. Introduction

While most children with ALL achieve complete remission with current therapeutic regimens, approximately 20% to 30% of them relapse and succumb to resistant disease. Various cellular mechanisms can give rise to multidrug resistance (MDR), but best studied is the transmembrane protein-mediated efflux of cytotoxic compounds which leads to a decreased intracellular drug accumulation and toxicity. P-glycoprotein (P-gp) is the best-characterized MDR pump.\textsuperscript{164} In addition, several other drug resistance-related proteins have been identified, including multidrug resistance-associated protein 1 (MRP1)\textsuperscript{165}, breast cancer resistance protein (BCRP)\textsuperscript{166} and lung resistance protein (LRP)\textsuperscript{167}.

Despite profound research, the clinical relevance of cellular resistance, mediated by MDR-related transport proteins, remains unclear. The comparison of data is hampered by the lack in standardized detection techniques. In addition, most assays used for clinical diagnosis are not sensitive enough to detect low-level resistant cells. Consequently, highly sensitive and specific tests are urgently needed in order to guarantee the accurate measurement of weak activity of MDR-related proteins.
2. Clinical relevance of multidrug resistance in childhood ALL

2.1. Paper 1

THE PROGNOSTIC SIGNIFICANCE OF MULTIDRUG RESISTANCE-RELATED PROTEINS IN CHILDHOOD ACUTE LYMPHOBLASTIC LEUKEMIA

Swerts K., De Moerloose B., Dhooge C., Laureys G., Benoit Y. and Philippé J.

European Journal of Cancer (Submitted)
The prognostic significance of multidrug resistance-related proteins in childhood acute lymphoblastic leukemia

Swerts K., De Moerloose B., Dhooge C., Laureys G., Benoit Y. and Philippé J.

European Journal of Cancer (Submitted)

ABSTRACT

An important problem in the treatment of children with acute lymphoblastic leukemia (ALL) is pre-existent or acquired resistance to structurally and functionally unrelated chemotherapeutic compounds. Various cellular mechanisms can give rise to multidrug resistance (MDR). Best studied is the transmembrane protein-mediated efflux of cytotoxic compounds which leads to decreased cellular drug accumulation and toxicity. Several MDR-related efflux pumps have been characterized, including P-glycoprotein (P-gp), multidrug resistance-associated protein 1 (MRP1), breast cancer resistance protein (BCRP) and lung resistance protein (LRP). P-gp expression and/or activity has been associated with unfavorable outcome in pediatric ALL patients whereas MRP1 does not seem to play a major role. BCRP and LRP might contribute to drug resistance in B-lineage ALL but larger studies are needed to confirm these results.

The present review summarizes the current knowledge concerning multidrug resistance-related proteins and focuses on the clinical relevance and prognostic value of these efflux pumps in childhood ALL.
INTRODUCTION

The prognosis of children diagnosed with acute lymphoblastic leukemia (ALL) has improved markedly during the past decades. However, approximately 25% of the affected children relapse and cannot be cured with current chemotherapy [1]. Intrinsic or acquired resistance to a wide variety of structurally and functionally unrelated chemotherapeutic compounds is one of the most important causes of treatment failure in childhood ALL. A variety of cellular mechanisms can give rise to multidrug resistance (MDR), including enhanced expression of cellular transporters, reduced drug uptake, alterations in detoxifying mechanisms, enhanced DNA repair processes, downregulation of drug targets, changes in cell cycle regulation and alterations in apoptotic pathways [2,3].

Classical multidrug resistance is associated with transmembrane protein-mediated efflux of cytotoxic compounds leading to a decreased cellular drug accumulation and toxicity. Several MDR-related drug efflux pumps have been characterized. Most of them belong to the superfamily of ATP-binding cassette (ABC) transporters.

P-glycoprotein (P-gp) is the best-characterized MDR pump [4]. Various neoplastic agents such as anthracyclines, mitoxantrone, taxanes, vinca alkaloids and epipodophyllotoxins are P-gp substrates. During the past 10 years, several new drug resistance-related proteins have been identified. One of these is the multidrug resistance-associated protein 1 (MRP1), an ATP-dependent efflux pump which extrudes glutathione conjugated compounds out of the cell [5]. Currently, the MRP family consists of nine MRP homologues (MRP1-MRP9). The breast cancer resistance protein (BCRP) is evolutionary distinct from the other ABC transporters [6]. BCRP is a half-molecule and formation of homo- or heterodimers is essential for its function as an active transporter [7,8]. Anthracyclines, topoisomerase I inhibitors, topoisomerase II inhibitors, cell-cycle inhibitors and methotrexate are BCRP substrates. Another MDR-related protein named lung resistance protein (LRP) is involved in the nuclear-cytoplasmic transport and/or sequestration of cytotoxic compounds. LRP is identified as the major vault protein (MVP) and confers resistance against vincristine, doxorubicin and etoposide [9]. Unlike P-gp, MRP and BCRP, LRP is not a member of the ABC transporter family. In the present review, current knowledge concerning the clinical relevance and prognostic value of multidrug resistance-related proteins in childhood ALL will be summarized.
**P-GLYCOPROTEIN**

Juliano and Ling were the first to isolate P-glycoprotein (P-gp) from resistant Chinese-hamster ovary cells [4]. P-gp is encoded by the MDR1 gene, located on the long arm of chromosome 7 (7q21.1) [10,11]. It is a 170 kDa protein which consists of two structurally homologous domains, each containing six hydrophobic transmembrane segments and a highly conserved ATP binding site [12,13]. The two times two domain organization is most likely the result of an internal gene duplication [14]. The presence of two ATP binding sites defines P-gp as a member of the ATP binding cassette (ABC) superfamily of transport proteins [15].

The mechanism by which P-gp decreases the intracellular accumulation of anthracyclines, mitoxantrone, taxanes, epipodophyllotoxins and vinca alkaloids, is poorly understood. There are at least three possible mechanisms of action. The first hypothesis assumes that amphiphatic and lipophilic substrates are removed from the cytoplasm through a channel formed by the transmembrane segments [16,17]. Furthermore, a “vacuum cleaner” model has been suggested in which P-gp transports compounds from either the inner or the outer leaflet of the lipid bilayer into the external medium [18,19]. Alternatively, P-gp might function as a “flippase”, transporting drugs from the inner to the outer leaflet of the bilayer after which the compounds will leave the plasma membrane by diffusion [20,21].

P-gp is expressed in various normal tissues with secretory or barrier functions including lung, placenta, testes, adrenal gland, kidney, liver, pancreas, colon, jejunum and brain [22-24]. In addition, P-gp is expressed by hematopoietic precursors and lymphocytes [25,26]. These findings suggest that P-gp plays a major role in the excretion and/or transport of cytotoxic xenobiotics.

P-gp mediated multidrug resistance can be reversed by various inhibitors. Competitive binding experiments showed that most modulators compete with cytotoxic substrates for P-gp binding sites [27]. Such agents include calcium channel blockers (e.g. verapamil), calmodulin inhibitors (e.g. pimozide), immunosuppressive agents (e.g. cyclosporine A, PSC 833), quinolones (e.g. chloroquine and quinine), indole alkaloids (e.g. reserpine), antibiotics (e.g. erythromycin), detergents (e.g. cremophor EL), steroids and anti-estrogens (e.g. tamoxifen) [28,29]. Other agents such as MDR1 anti-sense oligonucleotides, interference RNA and protein kinase C inhibitors (e.g. staurosporine), modulate P-gp activity through the transcriptional regulation of the MDR1 gene [30-32]. In addition, P-gp specific monoclonal antibodies such as MRK16, UIC2 and HYB-241 can be used to inhibit the P-gp mediated drug efflux [33-35]. Thus far, clinical trials with P-gp antagonists have yielded disappointing results [36-38].
Various techniques have been developed to study P-gp expression and function. MDR1 mRNA can be detected and/or quantified by reverse transcriptase PCR (RT-PCR) [39]. P-gp expression can also be determined using different monoclonal antibodies. Some recognize internal P-gp epitopes (e.g. C219 and JSB-1), others interact with external antigens (e.g. MRK16, 4E3 and UIC2) [40,41]. Monoclonal antibodies binding to external epitopes are preferred because they are superior in detecting low and variable levels of P-gp [42]. The P-gp activity is usually studied using functional flow cytometric assays. The intracellular accumulation of P-gp substrates such as rhodamine 123, JC-1, 3,3'-diethyloxacarbocyanine iodide (DiOC₂) or daunorubicin, is measured in the presence or absence of a P-gp specific inhibitor [43-45].

Several contradictory reports about the clinical importance of P-gp in childhood acute leukemia have been published [39,46-78]. Table 1 summarizes the reported data. Approximately 30% of all diagnostic ALL samples are P-gp positive. Some authors reported more P-gp positive patients at relapse [49,53,58,62] whereas others were unable to find a significant difference in the number of P-gp positive samples taken at initial diagnosis or at relapse [57,68]. When P-gp expression levels in initial ALL samples were compared to those in relapse samples, several authors reported a higher P-gp expression at relapse, i.e. especially at multiple relapses [48,55,56]. However, these results were not confirmed by others [39,46,61].

Tafuri et al. found an increased P-gp activity in samples from relapsed ALL patients [56]. However, Pieters and coworkers evaluated the daunorubicin and vincristin accumulation in the presence or absence of a P-gp inhibitor and found no difference in P-gp function between initial and relapse samples [50].

The P-gp expression did not differ between infant or older ALL patients [75,76]. Some authors found a significantly higher P-gp expression and/or function in T-ALL compared to B-ALL samples [62,73].

In some studies no relationship between P-gp status and event free survival (EFS) or overall survival (OS) was found [63,69,70,77,78]. However, most authors were able to demonstrate a significant association between P-gp expression and/or function and clinical outcome in childhood ALL. Brophy et al., Tafuri et al. and Kanerva et al. investigated the relationship between the response rate and P-gp expression [39,56,62]. Brophy and Tafuri reported a correlation between clinical response and absence of P-gp expression. By contrast, Kanerva did not find a relation between early response and P-gp status.

Goasguen and coworkers found no difference in the first complete remission rate between P-gp positive or P-gp negative patients [52]. However, relapses occurred more frequently in children with P-gp positive blasts and the survival rate was significantly higher in P-gp negative patients. The EFS-curve followed the same trend but did not reach statistical
significance. In accordance with Goasguen, Sauerbrey et al. found a significantly lower probability of remaining in continuous complete remission and a tendency for an increased relapse rate in P-gp positive patients [54]. Volm and coworkers published that P-gp positive patients had a significant lower median relapse-free survival [60]. In a study published by Dhooge and coworkers, a significant correlation between EFS and P-gp status was found [65]. The OS-curve followed the same trend but reached borderline statistical significance. These results were confirmed by Casale et al. [74]. De Moerloose et al. reported that the combination of P-gp expression and activity was a statistically significant parameter predicting relapse in childhood ALL [72]. Plasschaert and coworkers evaluated the P-gp expression and activity in 69 samples from pediatric and adult ALL patients and found an association between P-gp activity and EFS or OS in T-ALL patients [73]. By contrast, no impact of P-gp activity on EFS or OS was found in B-ALL patients.

In conclusion, the clinical importance of P-gp in childhood ALL remains unclear. One of the reasons for the variability in published data is the lack in standardized detection techniques. Moreover, the variety in detection methods hampers the comparison of data. Although this problem has been studied by a number of international workshops, implementation of consensus recommendations has been difficult [79-83]. Large and well controlled clinical studies, using standardized detection techniques, are indispensable to determine the prognostic significance of P-gp in childhood ALL.

MULTIDRUG RESISTANCE-ASSOCIATED PROTEIN 1

The multidrug resistance-associated protein (MRP1), which Cole et al. isolated from the doxorubicin resistant small cell lung cancer cell line H69AR, is also a member of the ATP-binding cassette transporter protein family [5]. The MRP1 gene, mapped on chromosome 16 (16p13.1), encodes a 190 kDa N-glycosylated hydrophobic anion pump localized on both the plasma and intracytoplasmic membranes, including the endoplasmic reticulum and Golgi apparatus [84].

The amino acid homology between MRP1 and P-gp is 15 %. Compared to P-gp, MRP1 has an additional N-terminal membrane spanning domain which is linked via a cytoplasmic loop to a P-gp-like core [85].
### Table 1: Clinical importance of P-gp in childhood ALL

(DOX: doxorubicin; DNR: daunorubicin; VCR: vincristin; FC: flow cytometry; I ALL: Initial ALL; R ALL: Relapsed ALL; IC: immunocytochemistry; RT-PCR: reverse transcription polymerase chain reaction; ISH: in situ hybridization; Acc: accumulation; Rho123: rhodamine 123; Mod: modulator; EFS: event-free survival; OS: overall survival; c: children; a: adults; WBC: white blood cell; CCR: continuous complete remission)

<table>
<thead>
<tr>
<th>Author</th>
<th>Method</th>
<th>Samples</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uckun et al., 1989 [46]</td>
<td>MDR1 gene expression: RNA blot  P-gp function: DOX Acc by FC</td>
<td>5 I ALL 5 R ALL</td>
<td>MDR1 gene expression level and intracellular DOX concentration did not differ between initial and relapse samples</td>
</tr>
<tr>
<td>Mizuno et al., 1991 [47]</td>
<td>P-gp expression: FC</td>
<td>9 I ALL 4 R ALL</td>
<td>P-gp was not detected in any of the samples</td>
</tr>
<tr>
<td>Geha et al., 1992 [48]</td>
<td>MDR1 gene expression: RNA slot blot</td>
<td>7 I ALL 14 1st R ALL 11 2nd R ALL</td>
<td>No difference in MDR1 mRNA levels between initial and relapsed ALL Significant increase in MDR1 expression levels in multiple relapsed ALL</td>
</tr>
<tr>
<td>Kingseman et al., 1992 [49]</td>
<td>P-gp expression: IC</td>
<td>48 I ALL 47 R ALL</td>
<td>At Dx: 6% P-gp expression; At relapse: 23% P-gp expression No relation between P-gp expression and clinical outcome in the relapsed patients</td>
</tr>
<tr>
<td>Pieters et al., 1992 [50]</td>
<td>P-gp expression: IC</td>
<td>28 I ALL 14 R ALL</td>
<td>P-gp was not detected in any of the tested samples DNR and VCR accumulation did not differ between resistant and sensitive cells Resistance modifiers did not enhance the in vitro cytotoxicity of DNR or VCR</td>
</tr>
<tr>
<td>Srinivas et al., 1997 [54]</td>
<td>P-gp expression: IC</td>
<td>36 I ALL 16 I ALL a</td>
<td>At Dx: 35% P-gp expression The rate of first complete remission did not differ between P-gp positive and P-gp negative patients Survival rate was significantly higher in P-gp negative compared to P-gp positive patients EFS curve followed this trend</td>
</tr>
<tr>
<td>den Boer et al., 1998 [55]</td>
<td>MDR1 gene expression: Cell lines from: RT-PCR and slot blot  P-gp function: IC and ISH</td>
<td>16 I ALL 20 R ALL</td>
<td>No difference in P-gp expression level between initial and relapsed ALL Complete clinical response to therapy correlated with absence of MDR1 expression</td>
</tr>
<tr>
<td>Goasguen et al., 1993 [52]</td>
<td>P-gp expression: IC</td>
<td>36 I ALL</td>
<td>At Dx: 35% P-gp expression The rate of first complete remission did not differ between P-gp positive and P-gp negative patients Survival rate was significantly higher in P-gp negative compared to P-gp positive patients EFS curve followed this trend</td>
</tr>
<tr>
<td>Beck et al., 1995 [55]</td>
<td>MDR1 gene expression: RT-PCR</td>
<td>27 I ALL 18 1st R ALL 7 2nd R ALL</td>
<td>No difference in MDR1 mRNA levels between initial and relapsed ALL Significant increase in MDR1 expression levels in multiple relapses</td>
</tr>
<tr>
<td>Pfaff et al., 1995 [56]</td>
<td>P-gp expression: FC</td>
<td>19 I ALL 14 R ALL</td>
<td>Significant difference in P-gp expression levels and activity between initial and relapsed patients Patients in complete remission showed a significant lower P-gp expression and function compared to those who failed to respond</td>
</tr>
<tr>
<td>Ivy et al., 1996 [57]</td>
<td>P-gp function: Rho123 + Mod by FC</td>
<td>30 I ALL 38 R ALL 22 I AML 8 R AML 5 Others</td>
<td>No difference in number of patients with P-gp expression at initial Dx or at relapse Significant increase in P-gp function at relapse</td>
</tr>
<tr>
<td>Volm et al., 1996 [58]</td>
<td>P-gp expression: IC</td>
<td>111 I ALL 28 R ALL</td>
<td>At Dx: 35% P-gp expression; At relapse: 54% P-gp expression At Dx, P-gp was more frequently expressed in patients who relapsed under therapy</td>
</tr>
<tr>
<td>Srinivas et al., 1997 [59]</td>
<td>P-gp expression: IC</td>
<td>88 I ALL</td>
<td>At Dx: 68% P-gp expression</td>
</tr>
<tr>
<td>Volm et al., 1997 [60]</td>
<td>P-gp expression: IC</td>
<td>104 I ALL</td>
<td>At Dx: 36% P-gp expression Patients with P-gp expression had a significant longer median relapse-free interval</td>
</tr>
<tr>
<td>den Boer et al., 1998 [61]</td>
<td>P-gp expression: FC</td>
<td>112 I ALL 7 2nd R ALL 20 I AML 6 1st R ALL 1 2nd R ALL</td>
<td>No difference in P-gp expression levels in initial or relapsed patients No significant correlation between P-gp expression and in vitro drug resistance P-gp expression did not relate to age, initial WBC count or unfavorable immunophenotype P-gp expression did not differ between AML and ALL patients</td>
</tr>
<tr>
<td>Karun et al., 1998 [62]</td>
<td>P-gp expression: FC</td>
<td>103 I ALL 15 R ALL</td>
<td>At Dx: 53% P-gp expression; At relapse: 73% P-gp expression No correlation between P-gp expression and early response at day 15 No relation between P-gp expression and age, sex, initial WBC count, FAB type or karyotype changes P-gp expression was significantly higher in T-ALL compared to B-ALL patients</td>
</tr>
<tr>
<td>Gurbuzar et al., 1998 [63]</td>
<td>MDR1 gene expression: RT-PCR</td>
<td>32 I ALL</td>
<td>No correlation between MDR1 gene expression levels and treatment outcome</td>
</tr>
<tr>
<td>den Boer et al., 1999 [64]</td>
<td>P-gp expression: FC</td>
<td>60 I ALL 25 R ALL</td>
<td>After correction for cell volume: Intracellular DNR concentration was lower in relapsed compared with initial ALL patients Inverse correlation between intracellular DNR concentration and in vitro DNR resistance No relation between P-gp expression and intracellular DNR concentration or in vitro DNR resistance</td>
</tr>
</tbody>
</table>

---

**Legend**

- **MTT assay**: Methyl thiazolyl tetrazolium assay
- **DOX Acc by FC**: Doxorubicin accumulation by flow cytometry
- **DNR acc and efflux**: Daunorubicin accumulation and efflux
- **Rho123 + Mod by FC**: Rhodamine 123 + modulator by flow cytometry
### Table 1: Clinical importance of P-pg in childhood ALL (Continued)

<table>
<thead>
<tr>
<th>Author</th>
<th>Method</th>
<th>Samples</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dhooge et al., 1999 [65]</td>
<td>P-gp expression: IC</td>
<td>102 I ALL</td>
<td>At Dx: 14% P-gp expression; At relapse: 34% P-gp expression P-gp positive patients at relapse had a 1.3-fold greater risk for adverse clinical outcome EFS was significantly higher in patients without P-gp expression OS curve followed the same trend but did not reach statistical significance P-gp expression was independent of WBC count, age, sex and karyotype P-gp expression was not induced by exposure to chemotherapy</td>
</tr>
<tr>
<td>Kikihara et al., 1999 [66]</td>
<td>MDR1 gene expression: RT-PCR</td>
<td>40 I ALL</td>
<td>No correlation between MDR1 gene expression and age or WBC count No difference in MDR1 expression levels between high- and low-risk patients</td>
</tr>
<tr>
<td>Ognatelli et al., 2000 [67]</td>
<td>MDR1 gene expression: RT-PCR</td>
<td>12 I ALL</td>
<td>MDR1 mRNA was undetectable in tested samples</td>
</tr>
<tr>
<td>Gurbuxani et al., 2000 [68]</td>
<td>MDR1 gene expression: RT-PCR</td>
<td>80 I ALL</td>
<td>At Dx: 16% MDR1 gene expression; MDR1 was not more frequently expressed at relapse Mean MDR1 mRNA levels were significantly higher for patients not achieving complete remission EFS was significantly higher in patients without P-gp expression OS curve followed the same trend but did not reach statistical significance P-gp expression was independent of WBC count, age, sex and karyotype P-gp expression was not induced by exposure to chemotherapy</td>
</tr>
<tr>
<td>Wuchter et al., 2000 [69]</td>
<td>P-gp expression: FC</td>
<td>102 I ALL</td>
<td>At Dx: 10% P-gp function No difference in P-gp expression or function between T-ALL and B-ALL Neither P-gp function, nor P-gp expression correlated with response to induction chemotherapy or OS No significant difference in P-gp expression or function between patients who relapsed and those in CCR</td>
</tr>
<tr>
<td>Kanerva et al., 2001 [70]</td>
<td>P-gp expression: FC</td>
<td>103 I ALL</td>
<td>No association between P-gp expression and EFS and OS No correlation between P-gp expression and coexpression of myeloid antigens</td>
</tr>
<tr>
<td>Dhooge et al., 2002 [71]</td>
<td>P-gp expression: IC</td>
<td>102 I ALL</td>
<td>At Dx: 14% P-gp expression; At relapse: 35% P-gp expression P-gp expression at diagnosis was associated with an increased risk for relapse Association between P-gp expression and EFS P-gp positive patients at relapse had a 2.2-fold greater risk for adverse clinical outcome</td>
</tr>
<tr>
<td>De Moerloose et al., 2003 [72]</td>
<td>P-gp expression: IC</td>
<td>52 I ALL</td>
<td>At Dx: 46% P-gp expression and 23% P-gp function EFS and OS were significantly lower in patients with P-gp expression and function Combination of P-gp expression and function is a statistically significant parameter predicting relapse</td>
</tr>
<tr>
<td>Pleaschaert et al., 2003 [73]</td>
<td>MDR1 gene expression: RT-PCR</td>
<td>36 I ALL c Children + adults: 35 I ALL a T-ALL than in B-ALL No correlation between P-gp activity and age, hemoglobin or platelet count at Dx Negative correlation between P-gp activity and WBC count In T-ALL: association between P-gp activity and OS In B-ALL: no impact of P-gp activity on EFS or OS</td>
<td></td>
</tr>
<tr>
<td>Casale et al., 2004 [74]</td>
<td>P-gp expression: FC</td>
<td>85 I ALL</td>
<td>At Dx: 47% P-gp expression and 30% P-gp function EFS was significantly lower in the P-gp positive population EFS was independent of age, WBC count, immunophenotype, FAB subtype and prednisolone response</td>
</tr>
<tr>
<td>Ramakers-van Woerden et al., 2004 [75]</td>
<td>MDR1 gene expression: FC</td>
<td>469 I ALL</td>
<td>No difference in P-gp expression between infants and older common/pre B-ALL patients ALL patients with MLL rearrangements had a 1.3-fold higher P-gp expression than MLL negative cases</td>
</tr>
<tr>
<td>Stam et al., 2004 [76]</td>
<td>MDR1 gene expression: real time RT-PCR</td>
<td>26 I ALL</td>
<td>No difference in MDR1 gene expression in infants compared to older children MDR1 expression levels did not correlate with in vitro drug resistance</td>
</tr>
<tr>
<td>Swerts et al., 2004 [77]</td>
<td>P-gp expression: FC</td>
<td>191 I ALL</td>
<td>ALL + AML + Other: Reversible JC-1 efflux in 20% and reversible Rho123 efflux in 27% of all samples No association between P-gp expression or activity and clinical outcome</td>
</tr>
<tr>
<td>Valera et al., 2004 [78]</td>
<td>MDR1 gene expression: RT-PCR</td>
<td>30 I ALL</td>
<td>MDR1 gene expression did not correlate with age, WBC count, race, immunophenotype, FAB type, CNS infiltration, MRD on day 28, EFS and OS</td>
</tr>
</tbody>
</table>

The spectrum of resistance caused by MRP1 overexpression is very similar to that of P-gp [86]. In vitro, enforced MRP1 expression confers resistance to anthracyclines, vinca alkaloids, epipodophyllotoxins, camptothecins and methotrexate, but not to taxanes and mitoxantrone [86-89]. Various glutathione, glucuronate or sulfate conjugates such as cysteinyl leukotriene LTC4, an important mediator of inflammatory responses, are also transported by MRP1 [90,91]. In addition, Rigato et al. reported that the transport of unconjugated bilirubin is mediated by MRP1 [92].
Glutathione (GSH) plays an important role in the MRP1 transport process. However, the exact mechanism by which GSH participates in the MRP1-mediated efflux is unknown. Some compounds such as LTC4 do not require GSH to be transported by MRP1 [90]. Other substrates such as daunorubicin, vincristine and rhodamine are only transported in the presence of GSH [84,93]. Based on these findings, Salerno and coworkers proposed a new working model [94]. They assumed that MRP1 is composed of two interlocked wheels. The first wheel binds GSH and LTC4 and functions as the power unit which turns when a substrate is bound and energy is provided by the hydrolysis of ATP. The second wheel is inert and turns only when it is connected to the first wheel. In the presence of saturating amounts of GSH and daunorubicin, both wheels are able to turn and expel one molecule of GSH and one molecule of daunorubicin, respectively [95].

MRP1 is expressed in most tissues in the human body, especially in lung, testes, kidney, skeletal muscle, epithelial and hematopoietic cells [5,96]. MRP1 is also expressed in the endothelial cells which form the blood-brain barrier [97]. These findings suggest that MRP1 plays an important role in the elimination and sequestration of cytotoxic drugs, leading to decreased concentrations at their target sites.

MRP1 does not seem to play a major role in multidrug resistance in childhood ALL (Table 2). Beck and coworkers did not find a difference in MRP1 gene expression levels in samples taken at diagnosis or at first relapse [55]. However, the MRP1 gene expression increased significantly in multiple relapse samples. Den Boer et al. evaluated the MRP1 expression in peripheral blood or bone marrow samples from 141 children with ALL [61]. In addition, the in vitro cytotoxicity of daunorubicin, vincristine, etoposide, prednisolone and L-asparaginase was evaluated. The MRP1 expression did not differ between samples taken at diagnosis or at relapse and no relation between MRP1 positivity and in vitro resistance was found. The MRP1 expression did not correlate with a prognostically unfavorable immunophenotype, white blood cell count or age. These results were confirmed by others [66,67,98]. In several studies, no correlation between MRP1 expression at diagnosis and event-free survival was found [78,98]. Plasschaert et al. analyzed samples from 36 children and 35 adults with de novo ALL and concluded that MRP1 activity had no prognostic impact on OS and EFS [73].
Table 2: Clinical significance of MRP1 in childhood ALL

<table>
<thead>
<tr>
<th>Author</th>
<th>Method</th>
<th>Samples</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beck et al., 1995 [55]</td>
<td>MRP1 gene</td>
<td>27 I ALL</td>
<td>No difference in MRP1 mRNA levels between initial and relapsed ALL patients</td>
</tr>
<tr>
<td></td>
<td>expression:</td>
<td>18 1st R ALL</td>
<td>Significant increase in MRP1 expression levels in multiple relapses</td>
</tr>
<tr>
<td></td>
<td>RT-PCR</td>
<td>7 ≥ 2nd R ALL</td>
<td></td>
</tr>
<tr>
<td>de Boer et al., 1998</td>
<td>MRP1 expression:</td>
<td>112 I ALL</td>
<td>No difference in MRP1 gene expression levels between initial and relapsed ALL patients</td>
</tr>
<tr>
<td></td>
<td>FC</td>
<td>22 1st R ALL</td>
<td>No significant correlation between MRP1 expression and in vitro drug resistance</td>
</tr>
<tr>
<td></td>
<td>MTT assay</td>
<td>7 ≥ 2nd R ALL</td>
<td>MRP1 expression did not relate to age, initial WBC count or unfavorable immunophenotype</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20 I ALL</td>
<td>MRP1 expression did not differ between AML and ALL patients</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 1st R ALL</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 2nd R ALL</td>
<td></td>
</tr>
<tr>
<td>Kakihara et al., 1999</td>
<td>MRP1 gene</td>
<td>40 I ALL</td>
<td>No correlation between MRP1 gene expression and age or WBC count</td>
</tr>
<tr>
<td></td>
<td>expression:</td>
<td>RT-PCR</td>
<td>No difference in MRP1 gene expression between high- and low-risk patients</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12 I ALL</td>
<td>MRP1 was overexpressed in most children with pre-B-ALL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MRP1 mRNA was undetectable in most children with T-ALL</td>
</tr>
<tr>
<td>Ogretmen et al., 2000</td>
<td>MRP1 gene</td>
<td>12 I ALL</td>
<td>MRP1 was overexpressed in most children with pre-B-ALL</td>
</tr>
<tr>
<td></td>
<td>expression:</td>
<td>RT-PCR</td>
<td>No difference in MRP1 gene expression between high- and low-risk patients</td>
</tr>
<tr>
<td></td>
<td></td>
<td>58 I ALL</td>
<td>MRP1 overexpression was not associated with unfavorable outcome</td>
</tr>
<tr>
<td></td>
<td></td>
<td>28 R ALL</td>
<td>MRP1 expression was independent of age, sex, initial blast count, FAB-type and immunological subtype</td>
</tr>
<tr>
<td>Sauerbrey et al., 2002</td>
<td>MRP1 gene</td>
<td>36 I ALL</td>
<td>No difference in MRP1 gene expression levels between initial and relapsed ALL patients</td>
</tr>
<tr>
<td></td>
<td>expression:</td>
<td>RT-PCR</td>
<td>No difference in MRP1 gene expression between high- and low-risk patients</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35 I ALL</td>
<td>High MRP1 activity did not influence EFS or OS in ALL patients</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 I ALL</td>
<td>No correlation between MRP activity and age, hemoglobin or platelet counts at diagnosis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasschaert et al., 2003</td>
<td>MRP1 gene</td>
<td>36 I ALL</td>
<td>At Dx: 16.6% MRP1 overexpression</td>
</tr>
<tr>
<td></td>
<td>expression:</td>
<td>RT-PCR</td>
<td>MRP1 gene expression did not correlate with age, WBC count, race, immunophenotype, FAB-type, CNS infiltration, MRD on day 28, EFS and OS</td>
</tr>
</tbody>
</table>

BREAST CANCER RESISTANCE PROTEIN

Chen and coworkers detected a 95 kDa ABC transporter in the human breast cancer cell line MCF-7/AdrVp [99]. RNA fingerprinting led to the identification of a 2.4-kb mRNA encoding a 655 amino acid protein which was termed breast cancer resistance protein (BCRP) [100]. The transporter is also known as ABCG2, placental transporter or mitoxantrone resistance protein. BCRP is encoded by the ABCG2 gene which was mapped on chromosome 4 (4q22) [101]. The transporter is a half-molecule with a C-terminal transmembrane segment and a N-terminal ATP-binding site [7]. The formation of homo- or heterodimers bridged by disulfide bonds is essential for its function as an active transporter [8]. In vitro, high BCRP expression causes resistance to anthracyclines (e.g. doxorubicin and daunorubicin), topo-isomerase I inhibitors (e.g. topotecan), topo-isomerase II inhibitors (e.g. bisantrene, etoposide and mitoxantrone), cell-cycle inhibitors (e.g. flavopiridol) and antifolates (e.g. methotrexate) [7,102-104]. However, a mutation in a single amino acid can change the substrate specificity and thereby alter the drug resistance profile. Cell lines overexpressing BCRP with an arginine at position 482 are able to transport mitoxantrone but not rhodamine 123 or doxorubicin, whereas cells with threonine or glycine at position 482 extrude rhodamine 123, doxorubicin and mitoxantrone [105].
Multidrug Resistance

BCRP overexpression has been described in resistant ovary, breast, colon and gastric cancer, fibrosarcoma cell lines, placental tissue, liver canalicular membranes, ducts and lobules of the breast, endothelium of veins and capillaries, epithelium of colon and small intestine and bile canaliculi [106].

Since children with ALL are treated with BCRP substrates such as methotrexate and doxorubicin, BCRP overexpression could be responsible for MDR. A few studies investigated the role of BCRP in childhood ALL (Table 3). In a retrospective study, Sauerbrey et al. analyzed the BCRP expression by TaqMan real-time RT-PCR in samples from 67 ALL patients (47 initial stage and 20 relapses) [107]. Children with T-ALL showed a lower BCRP expression than patients with precursor B-ALL. No relationship between BCRP expression and age, sex, initial blast count, prednisolone response or bone marrow response on day 15 and 33 was found. BCRP expression levels at relapse were not significantly different from those at diagnosis and BCRP overexpression was not associated with unfavorable outcome. Stam et al. reported a lower BCRP expression in infants than in older ALL patients [76].

These results indicate that BCRP could play a minor role in drug resistance in precursor B-ALL patients. However, larger studies are needed to evaluate the prognostic impact of BCRP expression and activity in childhood ALL.

Table 3: Clinical importance of BCRP in childhood ALL
(I ALL: Initial ALL; R ALL: Relapsed ALL; RT-PCR: reverse transcription polymerase chain reaction)

<table>
<thead>
<tr>
<th>Author</th>
<th>Method</th>
<th>Samples</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sauerbrey et al., 2002</td>
<td>BCRP gene expression RT-PCR</td>
<td>47 I ALL, 20 R ALL</td>
<td>BCRP gene expression was lower in T-ALL than in precursor B-ALL patients</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No difference in BCRP gene expression levels between initial and relapsed ALL patients</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No relationship between BCRP gene expression and age, sex, initial blast count, prednisolone response or bone marrow response on day 15 and day 33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No association between BCRP overexpression and unfavorable outcome</td>
</tr>
<tr>
<td>Stam et al., 2004 [76]</td>
<td>BCRP gene expression RT-PCR</td>
<td>26 I ALL</td>
<td>Infants expressed 2.4-fold less BCRP mRNA than older children with ALL</td>
</tr>
</tbody>
</table>

LUNG RESISTANCE PROTEIN

The lung resistance protein (LRP) was initially identified in an anthracycline-resistant, non-small cell lung cancer cell line which lacked P-gp overexpression [108]. The LRP gene is located on chromosome 16 (16p11.2), close to the MRP1 and protein kinase C-β gene, and encodes a 110 kDa protein [109]. Based on the LRP amino acid sequence, no transmembrane fragments or ATP-binding sites, characteristic for ABC transporters, were identified. Therefore, LRP is not considered to be a member of the ABC transporter family. Screening of an expression library identified LRP as the major vault protein (MVP) [9]. Vaults are highly conserved ribonucleoprotein organelles that are found in all higher eukaryotes.
They are localized in cytoplasmic vesicles and nuclear membranes and form the transporter core of the nuclear pore complex. Vaults are composed of the major vault protein, vault poly (ADP-ribose) polymerase, telomerase-associated protein 1 and small untranslated RNA [110]. The functional role of vaults in MDR is still unclear but it was proposed that they act by transporting drugs away from their subcellular targets by mediating the extrusion of cytostatics from the nucleus and/or the sequestration of drugs into vesicles. 

By immunocytochemistry, LRP has been found to be widely distributed in normal human tissues [111]. LRP overexpression has been observed in epithelia of the bronchus and digestive tract as well as in keratinocytes, adrenal cortex and macrophages. These results suggest that vaults play a role in detoxification processes.

The enforced expression of LRP in an ovarian carcinoma cell line led to increased numbers of vault particles but failed to confer drug resistance to etoposide, doxorubicin and vincristine [9]. Siva and coworkers concluded that the upregulation of vaults may be necessary but not sufficient to give rise to multidrug resistance [112].

Information about the clinical relevance of LRP in childhood ALL is limited (Table 4). In most studies, no difference in LRP expression between initial and relapsed ALL patients is found [67,98,113]. However, LRP expression was significantly higher in multiple relapse samples compared to diagnostic or first relapse samples [61]. Furthermore, LRP expression levels do not seem to differ between risk groups identified by WBC count, sex or age [61,66,98]. Ogretmen et al. found a high LRP expression in most children with pre-B ALL [67]. By contrast, LRP expression was much lower in T-ALL patients. These results are in accordance with those published by den Boer et al. [61]. Ramakers-van Woerden et al. found a high LRP expression in pro-B-ALL patients [75].

Volm et al. reported a significant association between LRP expression and long term survival in 38 children with de novo ALL [113]. This was confirmed by Sauerbrey et al. who demonstrated that children with a high LRP expression at diagnosis exhibited a lower tendency of remaining in first clinical remission [98].

The relationship between LRP expression and in vitro resistance to daunorubicin, vincristine, etoposide and prednisolone was also studied [61]. The LRP expression was weakly but significantly related to the in vitro resistance to daunorubicin. In addition, LRP expression correlated inversely with the intracellular accumulated daunorubicin concentration [64]. These findings suggest that LRP might contribute to drug resistance in children with ALL.
Table 4: Clinical relevance of LRP in childhood ALL

<table>
<thead>
<tr>
<th>Author et al.</th>
<th>Year</th>
<th>Samples</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volm et al., 1997 [113]</td>
<td></td>
<td>LRP expression: IC</td>
<td>38 I ALL; 25 R ALL</td>
</tr>
<tr>
<td>den Boer et al., 1998 [61]</td>
<td></td>
<td>LRP expression: FC</td>
<td>112 I ALL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>In vitro drug resistance: MTT assay</td>
<td>22 1st R ALL; 7 ≥ 2nd R ALL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>26 I AML; 20 1st R ALL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 1st R ALL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 2nd R ALL</td>
</tr>
<tr>
<td>den Boer et al., 1999 [64]</td>
<td></td>
<td>LRP expression: FC</td>
<td>60 I ALL; 25 R ALL</td>
</tr>
<tr>
<td>Kakihara et al., 1999 [65]</td>
<td></td>
<td>LRP gene expression: RT-PCR</td>
<td>40 I ALL</td>
</tr>
<tr>
<td>Ognitman et al., 2000 [67]</td>
<td></td>
<td>LRP gene expression: RT-PCR</td>
<td>12 I ALL</td>
</tr>
<tr>
<td>Sauerbrei et al., 2002 [98]</td>
<td></td>
<td>LRP gene expression: RT-PCR</td>
<td>58 I ALL; 28 R ALL</td>
</tr>
<tr>
<td>Ramakers-van Woerden et al., 2004 [75]</td>
<td></td>
<td>LRP expression: FC</td>
<td>469 I ALL</td>
</tr>
</tbody>
</table>

OTHER MULTIDRUG RESISTANCE PROTEINS

During the past few years, several new members of the ABC transporter family have been identified, including eight new MRP1 homologues (MRP2 or cMOAT, MRP3, MRP4, MRP5, MRP6, MRP7, MRP8 or ABCC11 and MRP9 or ABCC12), the sister of P-gp (SP-gp) and the transporter associated with antigen processing (TAP) [114-116]. Plasschaert et al. evaluated the MRP2 and MRP3 mRNA expression in childhood ALL [73]. In respectively 89% and 24% of the samples, MRP2 or MRP3 transcripts were detected. Steinbach and coworkers assessed the clinical relevance of MRP2-MRP5 and SMRP, a splice variant of MRP5, in 103 children with previously untreated ALL [117]. All five genes were expressed with great variability. However, only MRP3 expression was associated with a significantly worse prognosis. The median MRP3 levels were 10-fold higher in T-ALL compared to precursor B-ALL patients and 4-fold higher in boys than in girls. These findings suggest than MRP3 overexpression might account for the poor prognosis of male and T-ALL patients.

The clinical importance of the other transport proteins in childhood ALL remains to be elucidated.
CONCLUSIONS

One of the most important causes of treatment failure in childhood ALL is the emergence of multidrug resistance. Various mechanisms can give rise to clinical drug resistance but best studied is the overexpression of transmembrane transport proteins such as P-gp, MRP1, BCRP and LRP. In the present review, we focused on the clinical relevance and prognostic significance of these MDR-related efflux pumps in childhood ALL. Several authors found a relationship between P-gp expression and/or function and clinical outcome in pediatric ALL patients. Other studies, however, contradicted these findings. Data on the prognostic significance of other MDR proteins are scarce. LRP and BCRP might contribute to drug resistance in B-lineage ALL but larger studies are needed to confirm these results. MRP1 does not seem to play a major role in MDR in childhood ALL and of all newly identified ABC transporters (e.g. MRP2-MRP9, sP-gp, TAP), only MRP3 expression was associated with a worse prognosis.

Despite profound research, the clinical importance of MDR-related proteins in childhood ALL remains controversial. The comparison of data is hampered by the lack in standardized detection techniques, the heterogeneity in patient groups (e.g. pooled data of ALL and AML, initial and relapse samples, adults and children) and differences in treatment protocols. However, detailed information on the clinical relevance of MDR-related efflux pumps is needed before the potential of transporter-specific modulators can be studied. So far, phase 3 clinical trials with first- and second-generation P-gp antagonists have yielded conflicting results. This may be explained by the functional redundancy between different drug resistance efflux pumps. Moreover, limitations in the design of early resistance reversal trials contribute to the disappointing results.

It is also important to keep in mind that multidrug resistance is a multifactorial process. In addition to transmembrane transport proteins, other resistance mechanisms such as alterations in detoxification processes, apoptosis, DNA repair, cell cycle progression and drug uptake, might contribute to clinical drug resistance. Tools such as oligonucleotide-based or cDNA-based microarrays are relevant methods to screen for multifactorial mechanisms since they allow the determination of the expression profile of many genes in a single hybridization experiment [118-120]. Microarrays may open new avenues for the diagnosis of MDR in clinical samples. In addition, they are a helpful tool for assessing the role of transmembrane transport proteins in childhood ALL and will help us to obtain a clear picture on how to optimize treatment schedules in leukemia.
ACKNOWLEDGEMENTS

Research funded by a Ph.D. grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen) and the Children’s Cancer Fund ‘Kinderkankerfonds’.

REFERENCES


Multidrug Resistance


117. Steinbach D, Wittig S, Cario G et al. The multidrug resistance-associated protein 3 (MRP3) is associated with a poor outcome in childhood ALL and may account for the worse prognosis in male patients and T-cell immunophenotype. Blood 2003, 102, 4493-4498.


2.2. Paper 2

THE COMBINED ANALYSIS OF P-GLYCOPROTEIN EXPRESSION AND ACTIVITY PREDICTS OUTCOME IN CHILDHOOD ACUTE LYMPHOBLASTIC LEUKEMIA

De Moerloose B., Swerts K., Benoit Y., Laureys G., Loeys T., Philippé J. and Dhooge C.

Pediatric Hematology and Oncology 2003; 20: 381-391
THE COMBINED ANALYSIS OF P-GLYCOPEPTIDE EXPRESSION AND ACTIVITY PREDICTS OUTCOME IN CHILDHOOD ACUTE LYMPHOBLASTIC LEUKEMIA

Barbara De Moerloose, Katrien Swerts, Yves Benoit, and Geneviève Laureys
Department of Pediatrics, Ghent University Hospital, Ghent, Belgium

Tom Loeys
Department of Applied Mathematics and Computer Science, Ghent University, Ghent, Belgium

Jan Philippé
Department of Clinical Chemistry, Microbiology and Immunology, Ghent University Hospital, Ghent, Belgium

Catharina Dhooge
Department of Pediatrics, Ghent University Hospital, Ghent, Belgium

The link between drug resistance and relapse was often suggested, but rarely demonstrated in long-range clinical studies. Since it is nowadays recommended to validate immunocytological results, the authors studied prospectively 52 acute lymphoblastic leukemia (ALL) patients with an immunocytological test and a functional flow cytometric test. The 4-year EFS and OS were 79.3% and 85.2%, respectively. Patients scoring positive in both tests had a significantly higher relapse rate and worse survival (log rank \( p = .007 \) and \( .047 \) for event-free survival and overall survival, respectively). Among the different prognostic variables evaluated, only the combination of P-gp expression and activity was a statistically significant parameter predicting relapse in childhood ALL.

Keywords: childhood acute lymphoblastic leukemia, flow cytometry, immunocytology, multidrug resistance, P-glycoprotein

Acute childhood leukemias account for 25–30% of all childhood malignancies. The predominant type of leukemia in childhood is the acute lymphoblastic leukemia (ALL) [1]. After chemotherapy was introduced, it was clear that childhood ALL could be divided into many prognostically distinct subtypes and the concept of risk-adapted therapy was introduced [2, 3]. In childhood ALL, the following prognostic factors are frequently used: white blood cell count (WBC) at diagnosis, response to treatment, age of the patient at diagnosis, immunophenotype of the lymphoblasts, chromosomal aberrations, and minimal residual disease [1, 3–5]. Despite risk-adapted therapy, relapse occurs in 25% of the patients.
In a previous prospective study, we demonstrated the prognostic significance of the multidrug resistance (MDR) protein P-glycoprotein (P-gp) in childhood ALL [6]. P-gp is a 170-kD transmembrane protein, capable of extruding a variety of lipophilic compounds, including amphipathic neoplastic agents, out of the cell. Consequently, the intracellular drug concentration decreases and the cell becomes resistant to the drug [7]. Different structurally and functionally unrelated chemotherapeutic agents frequently used in the treatment of ALL, such as anthracyclines, vinca alkaloids, and epipodophyllotoxins, are substrates of P-gp action.

P-gp detection in patient samples can be performed by various techniques, which limits the comparison of the results of different clinical studies. Therefore, different P-gp detection assays have been evaluated in several workshops [8, 9]. At the moment, it is recommended to validate immunocytochemical test results by complementary tests, such as functional flow cytometric assays or reverse transcriptase–polymerase chain reaction tests.

In our previous study, an immunocytochemical assay was used [6]. In the present prospective study in childhood ALL, we combined immunocytochemical test results with P-gp flow cytometric activities, to determine which test or test combination gives the best prediction of relapse.

**PATIENTS AND METHODS**

**Patient Characteristics**

From June 1996 to January 2000, 61 newly diagnosed ALL patients were admitted to the Department of Pediatric Hematology/Oncology of the Ghent University Hospital. Out of these, 52 patients were included in the study and their prognostic variables are listed in Table 1. Nine children were excluded from the study. No material was available for P-gp determinations in 8 patients and 1 child died because of toxicity during induction chemotherapy. The clinical and biological characteristics and the relapse rate of these excluded patients were not significantly different from the studied population.

The diagnosis was based on standard morphology, cytochemistry, and immunophenotyping of the leukemic blasts. All patients were treated according to EORTC-CLCG (European Organization for Research and Treatment of Cancer—Childhood Leukemia Cooperative Group) protocols (45 with protocol 58881 and 7 with protocol 58951), including vincristine and anthracyclines. Six children, including the 2 patients with t(9;22), underwent a bone marrow (BM) transplant. The mean follow-up time was 39 months (median 40 months).

**Leukemic Samples**

Leukemic cells were collected from BM and/or peripheral blood. For immunocytochemistry, smears of BM and/or peripheral blood were air-dried for
TABLE 1 Distribution of Prognostic Factors

<table>
<thead>
<tr>
<th>Prognostic factor</th>
<th>Number of patients (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>1–10 years</td>
<td>38 (73)</td>
</tr>
<tr>
<td>&lt;1 or &gt;10 years</td>
<td>14 (27)</td>
</tr>
<tr>
<td>WBC count</td>
<td></td>
</tr>
<tr>
<td>&lt;50 × 10^9/L</td>
<td>39 (75)</td>
</tr>
<tr>
<td>&gt;50 × 10^9/L</td>
<td>13 (25)</td>
</tr>
<tr>
<td>Cytogenetic analysis</td>
<td></td>
</tr>
<tr>
<td>Hyperdiploidy</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Unfavorable</td>
<td>5 (10)</td>
</tr>
<tr>
<td>Other</td>
<td>44 (85)</td>
</tr>
<tr>
<td>Not determined</td>
<td>2 (4)</td>
</tr>
<tr>
<td>Blasts at day 8</td>
<td></td>
</tr>
<tr>
<td>&lt;1 × 10^9/L</td>
<td>47 (90)</td>
</tr>
<tr>
<td>&gt;1 × 10^9/L</td>
<td>5 (10)</td>
</tr>
<tr>
<td>DNA content</td>
<td></td>
</tr>
<tr>
<td>&gt;1.16 and &lt;1.6</td>
<td>5 (10)</td>
</tr>
<tr>
<td>&lt;1.16 and &gt;1.6</td>
<td>46 (88)</td>
</tr>
<tr>
<td>Not determined</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>28 (54)</td>
</tr>
<tr>
<td>Male</td>
<td>24 (46)</td>
</tr>
<tr>
<td>Immunophenotype</td>
<td></td>
</tr>
<tr>
<td>B-cell lineage</td>
<td>44 (85)</td>
</tr>
<tr>
<td>T-cell lineage or biphenotypic</td>
<td>8 (15)</td>
</tr>
</tbody>
</table>

Note. Hyperdiploidy was defined as >50 chromosomes in the blasts. Unfavorable cytogenetic analysis included t(9;22), t(4;11), 11q23, and near-haploidy.

24 h at room temperature and stored unfixed at −20°C. For flow cytometry, mononuclear cells from BM and/or peripheral blood were isolated by gradient centrifugation on Ficoll-hypaque, kept at 4°C, and analyzed within 24 h after prelevation. All samples contained more than 80% leukemic cells, based on morphological and immunological criteria. Dead cells were excluded with propidium iodide staining.

Control Cell Lines

Samples of the well-characterized human myeloid leukemia cell line K562 and its P-gp-expressing resistant clone K562/VLB20 [10] were included in each immunocytochemical and flow cytometric experiment as negative and positive controls, respectively.

Monoclonal Antibodies

In the immunocytochemical assay, P-gp was detected by two monoclonal antibodies (MoAbs), which recognize different extracellular epitopes of P-gp, namely 4E3 (Dako, Carpinteria, CA USA) and MRK16 (Kamiya Biomedical, Thousand Oaks, CA, USA). The optimal concentration for 4E3 and MRK16 was found to be 10 and 5 μg/mL, respectively.

Immunocytochemical Technique

P-gp expression on the leukemic blasts was demonstrated using the alkaline phosphatase–anti-alkaline phosphatase (APAAP) immunocytochemical
technique as described previously [11]. Among the different fixatives tested, 0.0125% glutaraldehyde in phosphate-buffered saline (PBS, pH 7.2) was found to be the optimal fixative to preserve both antigen expression and cellular morphology. The slides were fixed for 30 s at 4°C, washed, and incubated with the MoAb in a humidifier at room temperature for 30 min. The APAAP (APAAP-Dakopatts, Glostrup, Denmark) staining procedure was performed as prescribed by the manufacturer [12]. Higher assay sensitivity was achieved by repeat incubations with the rabbit anti-mouse antibodies and APAAP complex for an additional 5 min. Subsequently, the slides were developed with Fast-Red TR Salt (Sigma, St Louis, MO, USA) and counterstained with gallocyanin. Isotypic matched controls and controls without primary antibody were simultaneously examined to exclude atypical or background staining. Slides were examined by 3 independent observers. In agreement with the consensus recommendations on P-gp detection, no arbitrary minimal cutoff points were used [8]. In each experiment, a mean of 300 blasts were counted. In slides with poor cellularity, at least 50 blasts were counted. A patient was defined positive for P-gp expression if a red staining pattern along the cell membrane was observed with one of the MoAbs in at least one blast.

**Flow Cytometric Detection of P-gp Activity**

According to Ludescher et al. [13], 10^5 cells were incubated for 1 h at 37°C with 200 ng/mL rhodamine 123 (Rho 123, Sigma), which is a P-gp substrate, in the absence or presence of 10 μM of the P-gp inhibitor verapamil (Knoll, Ludwigshafen, Germany). The intracellular Rho 123 accumulation was measured on a FACSort flow cytometer (Becton Dickinson, Erembodegem, Belgium) calibrated with QC Windows and Quantum 1000 beads (Flow Cytometry Standards, San Juan, PR, USA). At least 10,000 events were counted and analyzed with Cell Quest software (Becton Dickinson, Erembodegem, Belgium). Cells were washed twice with cold (4°C) HBSS without phenol red and resuspended. After an additional incubation for 1 h at 37°C with or without 10 μM verapamil, the remaining intracellular Rho 123 retention after the efflux period was measured. A test result was considered positive if the cells exhibited Rho 123 efflux and if the intracellular Rho 123 fluorescence enhanced in the presence of verapamil by at least 10%, resulting in a Rho 123 retention ratio (RR) of at least 1.10 [14, 15].

**Statistical Analysis**

A comparison of frequency distribution of the different prognostic variables according to the P-gp status was determined by Fisher exact tests. Cumulative event-free survival (EFS) and overall survival (OS) were estimated according the Kaplan–Meier method and the log rank test was applied to
evaluate the difference between the P-gp-negative and -positive groups. Relative impact of baseline variables on event (= relapse or death) was expressed as risk ratios (RR). Univariate analysis was performed according to the Cox proportional hazard modeling. $p$ values less than .05 were considered to be statistically significant.

RESULTS

Remission Status, EFS, and OS

After induction treatment, complete remission (CR) was achieved in all patients. During the study period, 10 patients relapsed and out of these, 6 children died. Three children suffered from progressive leukemia; the other 3 died because of toxicity after BM transplant. The EFS at 4 years was 79.3% (95% confidence interval: 67.4–93.4%) and the OS 85.2% (95% confidence interval: 74.4–97.6%).

Results of the P-gp Detection Assays

Using the immunocytochemical technique, 24/52 patients (46%) had a positive test result. The EFS and OS were not significantly different in the P-gp-negative group, compared to the P-gp-positive group (log rank $p = .612$ and .228 for EFS and OS, respectively).

P-gp activity was found in 12 patients (23%) and was not significantly associated with a higher relapse rate (log rank $p = .134$) or an increased risk of death (log rank $p = .107$).

Patients who scored positive in either one of the tests (30/52) did not fare worse than the patients with 2 negative test results (log rank $p = .769$ and .168 for EFS and OS, respectively). However, the 6 children with P-gp expression in the immunocytochemical assay and simultaneously a positive functional test had a significantly worse outcome. Three of these 6 patients relapsed, compared to 7 (15%) in the group of children scoring negative for both tests or positive for only one test. Two positive children died, compared to 4/46 (9%) negative children. The cumulative EFS and OS of these patients are shown in Figure 1 (log rank $p = .007$ and .047, respectively).

Evaluation of Prognostic Factors

All prognostic variables were independently distributed in the P-gp-negative and -positive patient groups (Table 2). In the patient group scoring positive for the functional assay (FT), a borderline significant dependence ($p = .045$) with gender was seen. In our study, only the combination of P-gp expression and activity was a statistically significant parameter predicting relapse in childhood ALL ($p = .021$, Table 3). Moreover, the risk
FIGURE 1 Kaplan-Meier event-free survival and overall survival of 52 children with ALL, according to P-gp status (P-gp positive if both the immunocytochemical assay and the functional assay are positive).

of death was 4.49-fold higher in the patient group with a positive result for both P-gp expression and activity, compared to the negative group (\( p = .084 \), Table 3).

DISCUSSION

The link between drug resistance and relapse has often been suggested, but not defined in clinical treatment settings. The biological markers of drug resistance, such as P-gp, require clinical testing to evaluate them as potential therapeutic target. In the present prospective study, we report the results of immunocytochemical and functional detection of the MDR protein P-gp in childhood ALL and their clinical relevance. Previously, we only used immunocytochemistry for P-gp detection and we clearly demonstrated the potential of this technique to predict relapse [6].

Despite numerous advantages of immunocytochemistry (such as discrimination of malignant and normal cells and the evaluation of P-gp staining in individual cells) [16, 17], the French workgroup on P-gp detection does not recommend immunocytochemistry, merely because of a lack in sensitivity and reproducibility [9]. Others promote flow cytometry as the most valuable tool presently available for the detection of MDR in acute leukemia [18]. The functional flow cytometric assays are, in contrast to immunological tests,
<table>
<thead>
<tr>
<th>Prognostic factor</th>
<th>IC neg</th>
<th>IC pos</th>
<th>FT neg</th>
<th>FT pos</th>
<th>IC or FT neg</th>
<th>IC or FT pos</th>
<th>IC and FT neg</th>
<th>IC and FT pos</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–10 years</td>
<td>19</td>
<td>19</td>
<td>30</td>
<td>8</td>
<td>15</td>
<td>23</td>
<td>34</td>
<td>4</td>
<td>.532</td>
</tr>
<tr>
<td>&lt;1 or &gt;10 years</td>
<td>9</td>
<td>5</td>
<td>10</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>12</td>
<td>2</td>
<td>.713</td>
</tr>
<tr>
<td>WBC count</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50 × 10^9/L</td>
<td>21</td>
<td>18</td>
<td>28</td>
<td>11</td>
<td>16</td>
<td>23</td>
<td>33</td>
<td>6</td>
<td>.540</td>
</tr>
<tr>
<td>&gt;50 × 10^9/L</td>
<td>7</td>
<td>6</td>
<td>12</td>
<td>1</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>.540</td>
<td>.655</td>
</tr>
<tr>
<td>Cytogenetic analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hyperdiploid/other</td>
<td>24</td>
<td>21</td>
<td>36</td>
<td>9</td>
<td>19</td>
<td>26</td>
<td>41</td>
<td>4</td>
<td>.253</td>
</tr>
<tr>
<td>Unfavorable</td>
<td>2</td>
<td>3</td>
<td>.661</td>
<td>3</td>
<td>2</td>
<td>.301</td>
<td>3</td>
<td>2</td>
<td>.103</td>
</tr>
<tr>
<td>Blasts at day 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 × 10^9/L</td>
<td>25</td>
<td>22</td>
<td>36</td>
<td>11</td>
<td>19</td>
<td>28</td>
<td>42</td>
<td>5</td>
<td>.317</td>
</tr>
<tr>
<td>&gt;1 × 10^9/L</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>.399</td>
<td>4</td>
<td>.473</td>
</tr>
<tr>
<td>DNA content</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;1.16 and &lt;1.6</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>.580</td>
</tr>
<tr>
<td>&lt;1.16 and &gt;1.6</td>
<td>25</td>
<td>21</td>
<td>36</td>
<td>10</td>
<td>20</td>
<td>26</td>
<td>41</td>
<td>5</td>
<td>.480</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>16</td>
<td>12</td>
<td>25</td>
<td>3</td>
<td>15</td>
<td>13</td>
<td>26</td>
<td>2</td>
<td>.781</td>
</tr>
<tr>
<td>Male</td>
<td>12</td>
<td>12</td>
<td>.781</td>
<td>15</td>
<td>.045</td>
<td>7</td>
<td>17</td>
<td>.096</td>
<td>.397</td>
</tr>
<tr>
<td>Immunophenotype</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B-cell lineage</td>
<td>25</td>
<td>19</td>
<td>35</td>
<td>9</td>
<td>20</td>
<td>24</td>
<td>40</td>
<td>4</td>
<td>.367</td>
</tr>
<tr>
<td>T-cell lineage or</td>
<td>3</td>
<td>5</td>
<td>.447</td>
<td>5</td>
<td>2</td>
<td>.442</td>
<td>6</td>
<td>2</td>
<td>.227</td>
</tr>
<tr>
<td>biphentypic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. IC: result of the immunocytochemical assay (P-gp positive if either one of the MoAbs is positive); FT: result of the functional assay (P-gp positive if retention index > 1.10); IC or FT: combination of both assays (P-gp positive if either the immunocytochemical assay or the functional assay is positive); IC and FT: combination of both assays (P-gp positive if both assays are positive).
<table>
<thead>
<tr>
<th>Prognostic factor</th>
<th>Number of relapses/number of patients</th>
<th>RR</th>
<th>Number of deaths/number of patients</th>
<th>RR</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1–10 years</td>
<td>6/38</td>
<td>1</td>
<td>3/38</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>&lt;1 or &gt;10 years</td>
<td>4/14</td>
<td>1.83</td>
<td>3/14</td>
<td>2.83</td>
</tr>
<tr>
<td>WBC count</td>
<td>&lt;50 x 10^9/L</td>
<td>9/39</td>
<td>1</td>
<td>5/39</td>
<td>1.83</td>
</tr>
<tr>
<td></td>
<td>&gt;50 x 10^9/L</td>
<td>1/13</td>
<td>0.299</td>
<td>1/13</td>
<td>0.651</td>
</tr>
<tr>
<td>Cytogenetic analysis</td>
<td>Hyperdiploid/other</td>
<td>8/45</td>
<td>1</td>
<td>5/45</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Unfavorable</td>
<td>2/5</td>
<td>2.97</td>
<td>1/5</td>
<td>1.94</td>
</tr>
<tr>
<td>Blasts at day 8</td>
<td>&lt;1 x 10^9/L</td>
<td>9/47</td>
<td>1</td>
<td>5/47</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>&gt;1 x 10^9/L</td>
<td>1/5</td>
<td>1.19</td>
<td>1/5</td>
<td>1.89</td>
</tr>
<tr>
<td>DNA content</td>
<td>&gt;1.16 and &lt;1.6</td>
<td>0/5</td>
<td>1</td>
<td>0/5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>&lt;1.16 and &gt;1.6</td>
<td>10/46</td>
<td>uncal</td>
<td>uncal</td>
<td>6/46</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
<td>7/28</td>
<td>1</td>
<td>4/28</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>3/24</td>
<td>0.466</td>
<td>2/24</td>
<td>0.532</td>
</tr>
<tr>
<td>Immunophenotype</td>
<td>B-cell lineage</td>
<td>9/44</td>
<td>1</td>
<td>6/44</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>T-cell lineage or biphenotypic</td>
<td>1/8</td>
<td>0.688</td>
<td>0/8</td>
<td>0.001</td>
</tr>
<tr>
<td>IC</td>
<td>Negative</td>
<td>5/28</td>
<td>1</td>
<td>2/28</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>5/24</td>
<td>1.22 (0.351–4.25)</td>
<td>4/24</td>
<td>2.48 (0.453–13.5)</td>
</tr>
<tr>
<td>FT</td>
<td>Negative</td>
<td>6/40</td>
<td>1</td>
<td>3/40</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>4/12</td>
<td>2.88 (0.808–10.3)</td>
<td>3/12</td>
<td>3.76 (0.757–18.7)</td>
</tr>
<tr>
<td>IC or FT</td>
<td>Negative</td>
<td>4/22</td>
<td>1</td>
<td>1/22</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>6/30</td>
<td>1.18 (0.332–4.22)</td>
<td>5/30</td>
<td>3.99 (0.465–34.2)</td>
</tr>
<tr>
<td>IC and FT</td>
<td>Negative</td>
<td>7/46</td>
<td>1</td>
<td>4/46</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>3/6</td>
<td>4.98 (1.27–19.5)</td>
<td>2/6</td>
<td>4.49 (0.819–24.6)</td>
</tr>
</tbody>
</table>

Note. RR: risk ratio (95% confidence interval); uncal: uncalculable; IC: result of the immunocytochemical assay (P-gp positive if either one of the MoAbs is positive); FT: result of the functional assay (P-gp positive if retention index >1.10); IC or FT: combination of both assays (P-gp positive if either the immunocytochemical assay or the functional assay is positive); IC and FT: combination of both assays (P-gp positive if both assays are positive).
independent of the use of MoAbs, and they can directly evaluate the effect of MDR modulators [19]. The sensitivity of the functional test using Rho 123 has been demonstrated in clinical samples [20].

Using the functional test, 23% of our patients scored positive at initial diagnosis. This is in agreement with Tafuri et al., who observed 22.4% positive functional test results in childhood ALL [21]. The immunocytochemical P-gp expression rates in the present study were higher than the P-gp activity values (46 and 23%, respectively). In our previous study [6], only 14% of ALL patients scored positive for P-gp at initial diagnosis. We also could not confirm the prognostic value of immunocytochemistry as single test in the present study. The discrepancies between the two studies might be caused by the use of different MoAbs: C219 and 4E3 in the former and MRK16 and 4E3 in the present study. MRK16 and 4E3 were chosen because the externally binding MoAbs are more sensitive in the detection of low and variable levels of P-gp and because C219 is not specific for the MDR1 gene product since it also reacts with the MDR3 gene product [8, 22].

Current reports on P-gp expression, functional activity, and outcome in childhood ALL are conflicting. Goasguen et al. used C219 and JSB1 in an immunohistochemical assay and demonstrated a higher relapse rate and a shorter median survival for P-gp-positive patients [23]. Sauerbrey et al. also used C219 and immunocytochemistry and found a lower EFS in the P-gp-positive group [24].

In other studies in childhood ALL, flow cytometric functional or immunological assays were performed [15, 25]. Kanerva et al. also performed a prospective study in childhood ALL [25]. They used the MoAb JSB1, which recognizes an internal epitope of P-gp, in an immunological flow cytometric test and found no difference in EFS or OS between children with low or increased P-gp expression. Other researchers prefer the functional flow cytometric assay [15]. P-gp activity was studied by Wuchter et al. in 102 children with ALL and did not correlate with immunophenotypic subgroups, response to induction chemotherapy, relapse rate, and OS [15].

These conflicting data resulting from prospective studies based on one detection technique stress the importance of combining complementary tests, such as immunocytochemistry and functional assays. Recently, Damiani et al. found a shorter disease-free survival in adults with ALL if P-gp overexpression was associated with a defect in daunorubicin accumulation [26].

Among the different variables studied in our 52 ALL patients, only the simultaneous positivity of P-gp activity and P-gp expression was significantly associated with a higher risk for relapse. No single test results and no other prognostic parameters in childhood ALL (age, gender, WBC count at diagnosis, number of blasts at day 8, immunophenotype of the blasts, DNA content, and cytogenetic abnormalities) had a comparable increased risk for relapse or death.
**CONCLUSION**

The combined use of a functional flow cytometric assay with Rho 123 and an immunocytochemical staining of P-gp not only allows one to determine accurately P-gp in clinical samples, as was suggested by others [27], but additionally might contribute to the identification of patients with increased risk for relapse.

**REFERENCES**


3. Evaluation of P-glycoprotein activity

3.1. Paper 3

COMPARISON OF TWO FUNCTIONAL FLOW CYTOMETRIC ASSAYS TO ASSESS P-GP ACTIVITY IN ACUTE LEUKEMIA

Swerts K., De Moerloose B., Dhooge C., Noens L., Laureys G., Benoit Y. and Philippé J.

Leukemia & Lymphoma 2004; 45: 2221-2228
Comparison of Two Functional Flow Cytometric Assays to Assess P-gp Activity in Acute Leukemia

KATRIEN SWERTS a,*, BARBARA DE MOERLOOSE a, CATHARINA DHOOGE a, LUCIEN NOENS b, GENEVIÈVE LAUREYS a, YVES BENIOIT a and JAN PHILIPPE b

aDepartment of Pediatric Hematology and Oncology, Ghent University Hospital, Ghent, Belgium; bDepartment of Hematology, Ghent University Hospital, Ghent, Belgium; cDepartment of Clinical Chemistry, Microbiology and Immunology, Ghent University Hospital, Ghent, Belgium

(Received 10 June 2004)

One of the possible causes of treatment failure in acute leukemia is the emergence of multidrug resistance caused by P-glycoprotein (P-gp) overexpression. We compared a flow cytometric assay using JC-1 with a technique using rhodamine 123 (rho123) to evaluate the P-gp function in acute leukemia. Samples from 50 acute leukemia patients were analyzed by both functional assays. The P-gp expression was assessed by an immunological flow cytometric test and the association between the P-gp status and the clinical outcome was evaluated. Of all samples, 28% showed a reversible JC-1 efflux and 36% scored positive for the rho123 assay. In two cases, the leukemic blasts showed a reversible JC-1 efflux whereas they were negative for rho123. These patients had blast cells with a very low P-gp activity. Six samples scored positive for the rho123 assay but were negative for the JC-1 test. Five of these samples did not express P-glycoprotein and were considered false positive. We found a strong correlation between the JC-1 and the rho123 test ($R_s = 0.59, p < 0.0001$) and the JC-1 and the immunological assay ($R_s = 0.29, P = 0.05$). There was also an association between the JC-1 status and the clinical outcome of adult patients ($\chi^2 = 6.30, P = 0.04$). In conclusion, we recommend the JC-1 assay to study the P-gp activity in acute leukemia because it is more specific and less labor intensive than conventional functional flow cytometric tests using rhodamine 123. In addition, the JC-1 assay can be used to identify adult patients with an increased risk for adverse clinical outcome.

Keywords: Multidrug resistance; P-glycoprotein; JC-1; Rhodamine 123; Acute leukemia; Flow cytometry

INTRODUCTION

Treatment results in acute leukemia have improved markedly during the past 30 years. However, some patients relapse and cannot be cured with current chemotherapy due to the emergence of multidrug resistance. Several mechanisms of resistance are known, but best studied is the overexpression of a 170 kDa transmembrane protein, P-glycoprotein (P-gp). This ATP-dependent efflux pump extrudes a variety of lipophylic, cytotoxic compounds out of the cell and hence decreases their intracellular concentration and toxicity [1]. Different structurally and functionally unrelated chemotherapeutic agents used in the treatment of acute leukemia such as anthracyclines, vinca alkaloids and epipodophyllotoxins, are substrates of the P-gp pump.

In acute myeloid leukemia, the prognostic role of P-gp has been clearly established [2–4]. However, the clinical value of P-gp in childhood acute leukemia remains unclear [5–10]. One of the reasons for the variability in the published data is the variety of detection techniques. In addition, there is a lack of sensitive and specific detection methods. These issues have been addressed in a number of international workshops and several recommendations have been formulated [11–14]. Especially, the use of a functional flow cytometric assay was promoted because it allows, in contrast to other tests, the detection of P-gp activity [13,14].

A recently developed flow cytometric assay evaluates the P-gp function by measuring the JC-1 accumulation in the absence or the presence of a P-gp inhibitor [15]. JC-1 is a carbocyanine liquid crystal forming probe which was initially used to analyze the mitochondrial potential [16].
Due to the stacking in a liquid crystal form, the fluorescence emission wavelength of this probe depends on its concentration. When the JC-1 monomers are excited at 488 nm, the emission spectrum reaches its maximum at 537 nm (green fluorescence). Beyond a critical concentration, JC-1 aggregates are formed. The aggregates display a fluorescent emission centered at 597 nm (red fluorescence) in addition to the green fluorescence. Sensitive cells display both green and red fluorescence. Resistant cells, expressing functional P-gp pumps will contain lower concentrations of JC-1. Hence they lose the red fluorescence and display a decrease in green fluorescence intensity.

In this study, we compared this new functional assay using JC-1 with a conventional technique using rhodamine 123 to evaluate the P-gp function in acute leukemia. In addition, we studied the relation between the P-gp function and expression in leukemic cells and evaluated the association between the P-gp status and the clinical outcome of the patients.

**MATERIALS AND METHODS**

**Patients and Sample Preparation**

Between May 2001 and October 2002, the P-gp function and expression in peripheral blood or bone marrow samples, taken at diagnosis or relapse, from 50 patients suffering from acute leukemia (acute lymphoblastic leukemia (ALL) as well as acute myeloid leukemia (AML)), were evaluated. We analyzed samples from 30 children and 20 adults. The patient characteristics are summarized in Table I. Approval from the Ethical Committee was obtained.

<table>
<thead>
<tr>
<th>TABLE I Patient characteristics.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pediatric Patients</td>
</tr>
<tr>
<td>precursor B-ALL at diagnosis</td>
</tr>
<tr>
<td>at relapse</td>
</tr>
<tr>
<td>T-ALL at diagnosis</td>
</tr>
<tr>
<td>Biphenotypic AL at diagnosis</td>
</tr>
<tr>
<td>AML at diagnosis</td>
</tr>
<tr>
<td>at relapse</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>BM samples</td>
</tr>
<tr>
<td>PB samples</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Adult Patients</td>
</tr>
<tr>
<td>precursor B-ALL at diagnosis</td>
</tr>
<tr>
<td>at relapse</td>
</tr>
<tr>
<td>Bilinear AL at diagnosis</td>
</tr>
<tr>
<td>AML at diagnosis</td>
</tr>
<tr>
<td>at relapse</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>BM samples</td>
</tr>
<tr>
<td>PB samples</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
</tr>
</tbody>
</table>

ALL: acute lymphoblastic leukemia; AL: acute leukemia; AML: acute myeloid leukemia; BM: bone marrow; PB: peripheral blood

**Control Cell Lines**

Samples from the human chronic myeloid leukemia cell line K562 and its P-gp expressing, resistant clone K562/VLB were included in each flow cytometric experiment as a negative and a positive control, respectively [17].

**Flow Cytometry**

Fluorescence was analyzed on a FACSort flow cytometer (BD Biosystems, Erembodegem, Belgium) calibrated with the SPHERO™ Calilow Kit (Spherotech, Libertyville, USA). The FACSort flow cytometer is equipped with an argon ion laser emitting a 488 nm beam and a red diode laser emitting a 635 nm beam which makes it possible to analyze 6 parameters (forward scatter, side scatter and 4 fluorescence parameters) at the same time. At least 10,000 events were counted and analyzed with the Cell Quest software (BD Biosystems, Erembodegem, Belgium).

**Identification of Leukemic Blasts**

Mononuclear cells (5.10^5) were incubated with a monoclonal antibody, labeled with allopbycocyanin, for 30 min at room temperature. The monoclonal antibodies CD34 (Clone 8G12), CD19 (Clone SJ25C1) or CD3 (Clone SK7) provided by BD Biosystems (BD Biosystems, Erembodegem, Belgium) or CD33 (Clone D3HL60.251, Beckman Coulter, Fullerton, USA) were used to identify the leukemic blasts. Debris, dead cells and hematogones were excluded using the forward and side scatter. We identified the leukemic cells combining the tumor specific monoclonal antibody with the forward and side scatter.

**Flow Cytometric Detection of P-gp Activity using JC-1**

After incubation with a tumor specific monoclonal antibody, samples were washed twice and cells were resuspended in 1 ml PBS (phosphate-buffered saline) containing 1 g/l glucose and 0.1 μM JC-1 monomers (Molecular Probes, Eugene, USA) with or without the P-gp inhibitor cyclosporine A (2 μM) (Novartis AG, Basel, Switzerland) [18]. The glucose was added to maintain the intracellular ATP level and the solution of the JC-1 monomers (0.1 μM) was prepared by diluting 1% (v/v) of the stock solution (1 g/l) in PBS. This solution was then filtered on a cellulose syringe filter.
Flow Cytometric Detection of P-gp Activity using Rhodamine 123

After incubation with a tumor specific monoclonal antibody, the cells were washed and incubated for one hour at 37°C with 200 ng/ml rhodamine 123 (Sigma, St. Louis, USA) with or without 10 μM of the P-gp inhibitor verapamil (Knoll AG, Ludwigshafen, Germany). The cells were washed twice and resuspended in medium without rho123 in the presence or absence of verapamil. After an incubation of 1 h at 37°C, the intracellular rho123 concentration was measured with a flow cytometer. The percentage of cells showing a rhodamine 123 efflux which could be inhibited by verapamil was determined.

The threshold for positivity of the rho123 assay was calculated using ROC curve analysis [19,20]. The cut-off value corresponding with the highest accuracy (percentage of positive cells = 0.49) was accepted as the threshold for positivity.

Flow Cytometric Detection of P-gp Expression

Mononuclear cells (5.10^5) isolated from peripheral blood or bone marrow were incubated for 30 min with the unlabeled isotypic control antibody IgG2a (Clone: DAK-G05, Dako Corporation, Glostrup, Denmark) or with the unlabeled P-gp specific monoclonal antibody MRK16 (Clone: MRK16, Kamiya Biomedical Company, Thousand Oaks, USA) in parallel samples. Cells were washed twice and incubated for 20 min with fluorescein isothiocyanate- or phycoerythrin- labeled polyclonal goat F(ab')2 anti-mouse IgG antibodies (Caltag Laboratories, Burlingame, UK). Leukemic blasts were identified by tumor specific markers and the fluorescence was analyzed on the flow cytometer.

Using the Kolmogorov-Smirnov test, the fluorescence of the gated leukemic cell population incubated with MRK16 was compared to the fluorescence of the cells incubated with the isotypic control antibody IgG2a. A D-value of 0.15 was used as a threshold for positivity [22].

Statistical Analysis

We used the Spearman’s rank correlation coefficient (R_s) to study the correlation between different parameters. After classifying the results in a frequency table, a χ^2 test

FIGURE 1 Example of a positive (A) and a negative (B) sample. The histograms show the changes in red fluorescence (FL2). Resistant cells, expressing functional P-gp pumps will contain lower concentrations of JC-1. Below a critical concentration, no JC-1 aggregates displaying a red fluorescence are formed. Therefore, positive cells show a shift in red fluorescence which can be reversed by cyclosporin A, a P-gp inhibitor. The percentage of leukemic blasts showing a reversible shift in red fluorescence was determined. The threshold for positivity was calculated using ROC analysis and was set at 18.7%.
was performed to study the relationship between the different categories. Finally, we compared proportions using a $\chi^2$ test with a Yates’ correction for continuity.

RESULTS

Analysis of the P-gp Activity using Rho123 and JC-1

In samples from 50 acute leukemia patients, the P-gp activity was analyzed using both functional assays. The results are summarized in Table II.

Of all tested samples, 28% (14/50) showed a reversible JC-1 efflux. 36% (18/50) of all samples scored positive for the functional assay using rho123. Twelve samples scored positive for both tests and thirty patients showed no functional P-gp pumps. In two pediatric cases, the leukemic blasts showed a reversible JC-1 efflux but they did not transport rho123. Six patients scored positive for the rho123 test but were negative for the JC-1 assay.

Significantly more AML than ALL patients had leukemic blasts which expressed functional P-gp pumps (for JC-1 assay: $\chi^2 = 5.53$, $P = 0.019$; for rho123 assay: $\chi^2 = 8.27$, $P = 0.004$). Eleven out of 24 AML patients (46%) scored positive for the functional JC-1 assay. Fourteen AML patients (54%) were positive for the rho123 test. In contrast, only 12% of the ALL patients (3/26) showed a reversible JC-1 efflux. Four ALL patients (15%) scored positive for the rho123 test.

The correlation between both functional tests was strong ($R_s = 0.59$, $p < 0.0001$). Based on the thresholds for positivity, the results were classified in a frequency table and the relationship between the JC-1 and the rho123 assay was calculated using a $\chi^2$ test. There was a clear association between both tests ($\chi^2 = 17.97$, $p < 0.0001$). There was also a correlation between the P-gp function evaluated by the JC-1 test and the P-gp expression analyzed using the monoclonal antibody MRK16 ($R_s = 0.29$ and $P = 0.05$). In contrast, there was no correlation between the rho123 assay and the immunological MRK16 test ($R_s = 0.12$ and $P = 0.43$).

Correlation between the Functional JC-1 Assay and the Clinical Outcome

The patients were divided into a pediatric and an adult group. The mean follow-up time was 40 months for the surviving adult leukemia patients ($n = 10$).

Out of 30 children, three died of leukemic disease (two AML patients and one precursor B-ALL patient) and one child with a biphenotypic acute leukemia relapsed. None of these four children showed a reversible JC-1 efflux (Table III). Only one scored positive for the rho123 assay. Six children scored positive for the JC-1 assay but were still in complete remission at the time of the analysis. There was no association between the P-gp activity and the clinical outcome in the pediatric patients (for JC-1 assay: $\chi^2 = 1.15$ and $P = 0.56$; for rho123 test: $\chi^2 = 0.43$ and $P = 0.81$). There was also no association between the results of the immunological assay and the clinical outcome ($\chi^2 = 0.56$ and $P = 0.76$).

From a child with AML, two samples, taken at different time points, were analyzed using the JC-1 assay. The first sample was taken at relapse and the leukemic blasts did not express functional P-gp pumps. The second sample was collected after 3 weeks of chemotherapy and it showed an increased percentage of leukemic blasts with a reversible shift in red fluorescence (Fig. 2). The child died shortly thereafter.

Out of 20 adult patients, 10 (50%) died of leukemic disease and one patient relapsed (Table III). Eight patients (40%) showed a reversible JC-1 efflux. One of them was still in complete remission at the time of the analysis. There was an association between the JC-1 status and the clinical outcome of the adult acute leukemia patients ($\chi^2 = 6.30$ and $P = 0.04$). Ten adults scored positive for the functional rho123 test. Of these, 3 were still in complete remission at the time of analysis. There was no association between the results of the rho123 or the immunological test and the clinical outcome of the adults (for rho123: $\chi^2 = 2.40$ and $P = 0.30$; for MRK16: $\chi^2 = 3.24$ and $P = 0.20$).

Based on the results of the adult patients, the sensitivity, the specificity, the positive predictive value and the negative predictive value of the functional assay using JC-1 was calculated. They were 64, 89, 88 and 67%, respectively.

DISCUSSION

Rhodamine 123 is most frequently used as substrate in functional flow cytometric assays evaluating P-gp function. It has a good fluorescence quantum yield but a fraction of the fluorophore binds to the cell membrane [23]. This leads to a high fluorescence background and compensation problems which hamper the identification of leukemic blasts. To overcome this problem, we incubated the blasts with a tumor specific monoclonal antibody labeled with allophycocyanin. This fluorochrome can only be excited at 600–640 nm. Using a flow cytometer equipped with two lasers, allophycocyanin can be excited without causing a high fluorescence background or compensation problems related to rhodamine 123. This makes it possible to identify small populations of leukemic blasts or detect leukemic cells with a low P-gp activity. Moreover, better results are obtained with a rho123 assay based on efflux rather than uptake [18]. However, efflux procedures are time consuming. Finally, rho123 is also a substrate for MRPI although it is less efficiently transported by MRPI than by P-gp [24].

JC-1 causes no background problems because all the detected red fluorescence results from the liquid crystals in the mitochondria of the cells [25]. In addition, this test...
does not require drug efflux monitoring. This makes the assay much faster and less labor intensive compared to the conventional functional flow cytometric assay using rho123. In contrast with the rho123 assay which takes more than 3 h, the JC-1 test is completed in less than 1 h. Finally, Legrand et al. found that JC-1 could not be transported by MRP1 which increases the specificity of the test [15].

TABLE II Results of the flow cytometric evaluation of the P-gp activity using JC-1 and rho123. Samples 1 to 30: pediatric patients; Samples 31 to 50: adult patients.

<table>
<thead>
<tr>
<th>NR</th>
<th>TYPE</th>
<th>Time Point</th>
<th>Sample</th>
<th>% JC-1 positive cells</th>
<th>JC-1</th>
<th>% rho123 positive cells</th>
<th>rho123</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>18,72</td>
<td>+</td>
<td>8,96</td>
<td>+</td>
<td>CR</td>
</tr>
<tr>
<td>2</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>12,44</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>3</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>1,56</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>4</td>
<td>AML</td>
<td>Relapse</td>
<td>BM</td>
<td>18,26</td>
<td>-</td>
<td>14,29</td>
<td>+</td>
<td>CR</td>
</tr>
<tr>
<td>5</td>
<td>B-ALL</td>
<td>Relapse</td>
<td>BM</td>
<td>13,89</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>6</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>PB</td>
<td>59,78</td>
<td>+</td>
<td>3,11</td>
<td>+</td>
<td>CR</td>
</tr>
<tr>
<td>7</td>
<td>AML</td>
<td>Relapse</td>
<td>PB</td>
<td>6,19</td>
<td></td>
<td>2,36</td>
<td>+</td>
<td>D</td>
</tr>
<tr>
<td>8</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>3,65</td>
<td>-</td>
<td>0,37</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>9</td>
<td>B-ALL</td>
<td>Relapse</td>
<td>BM</td>
<td>0,01</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>D</td>
</tr>
<tr>
<td>10</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>PB</td>
<td>12,33</td>
<td>-</td>
<td>0,49</td>
<td>+</td>
<td>CR</td>
</tr>
<tr>
<td>11</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>PB</td>
<td>3,77</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>12</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>31,76</td>
<td>+</td>
<td>18,51</td>
<td>+</td>
<td>CR</td>
</tr>
<tr>
<td>13</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>13,60</td>
<td>-</td>
<td>0,16</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>14</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>11,78</td>
<td>-</td>
<td>1,81</td>
<td>+</td>
<td>CR</td>
</tr>
<tr>
<td>15</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>14,67</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>16</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>PB</td>
<td>37,80</td>
<td>+</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>17</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>1,91</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>18</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>12,39</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>19</td>
<td>T-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>0,44</td>
<td>-</td>
<td>0,23</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>20</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>4,24</td>
<td>-</td>
<td>0,30</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>21</td>
<td>Bi. ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>8,39</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>R</td>
</tr>
<tr>
<td>22</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>8,55</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>23</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>29,11</td>
<td>+</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>24</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>2,18</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>25</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>PB</td>
<td>13,61</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>26</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>6,28</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>27</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>53,60</td>
<td>+</td>
<td>17,09</td>
<td>+</td>
<td>CR</td>
</tr>
<tr>
<td>28</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>4,01</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>29</td>
<td>T-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>0,00</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>30</td>
<td>AML</td>
<td>Relapse</td>
<td>BM</td>
<td>4,24</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>D</td>
</tr>
<tr>
<td>31</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>0,63</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>D</td>
</tr>
<tr>
<td>32</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>5,94</td>
<td>-</td>
<td>4,05</td>
<td>+</td>
<td>CR</td>
</tr>
<tr>
<td>33</td>
<td>AML</td>
<td>Diagnosis</td>
<td>PB</td>
<td>18,81</td>
<td>+</td>
<td>10,98</td>
<td>+</td>
<td>D</td>
</tr>
<tr>
<td>34</td>
<td>AML</td>
<td>Diagnosis</td>
<td>PB</td>
<td>27,63</td>
<td>+</td>
<td>3,78</td>
<td>+</td>
<td>D</td>
</tr>
<tr>
<td>35</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>24,67</td>
<td>+</td>
<td>1,30</td>
<td>+</td>
<td>D</td>
</tr>
<tr>
<td>36</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>4,59</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>37</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>5,46</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>D</td>
</tr>
<tr>
<td>38</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>10,32</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>39</td>
<td>B-ALL</td>
<td>Relapse</td>
<td>BM</td>
<td>8,61</td>
<td>-</td>
<td>0,17</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>40</td>
<td>AML</td>
<td>Relapse</td>
<td>BM</td>
<td>57,66</td>
<td>+</td>
<td>4,00</td>
<td>+</td>
<td>D</td>
</tr>
<tr>
<td>41</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>11,03</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>D</td>
</tr>
<tr>
<td>42</td>
<td>AML</td>
<td>Diagnosis</td>
<td>PB</td>
<td>85,24</td>
<td>+</td>
<td>8,77</td>
<td>+</td>
<td>CR</td>
</tr>
<tr>
<td>43</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>BM</td>
<td>4,63</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>44</td>
<td>Bilinear</td>
<td>Diagnosis</td>
<td>PB</td>
<td>13,43</td>
<td>-</td>
<td>2,04</td>
<td>+</td>
<td>CR</td>
</tr>
<tr>
<td>45</td>
<td>B-ALL</td>
<td>Diagnosis</td>
<td>PB</td>
<td>14,13</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>46</td>
<td>AML</td>
<td>Diagnosis</td>
<td>PB</td>
<td>2,67</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>D</td>
</tr>
<tr>
<td>47</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>86,11</td>
<td>+</td>
<td>11,07</td>
<td>+</td>
<td>R</td>
</tr>
<tr>
<td>48</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>29,72</td>
<td>+</td>
<td>4,13</td>
<td>+</td>
<td>D</td>
</tr>
<tr>
<td>49</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>5,66</td>
<td>-</td>
<td>0,00</td>
<td>-</td>
<td>CR</td>
</tr>
<tr>
<td>50</td>
<td>AML</td>
<td>Diagnosis</td>
<td>BM</td>
<td>29,42</td>
<td>+</td>
<td>27,77</td>
<td>+</td>
<td>D</td>
</tr>
</tbody>
</table>

ALL: acute lymphoblastic leukemia; AML: acute myeloid leukemia; Bi.: biphenotypic; BM: bone marrow; PB: peripheral blood; D: deceased; R: relapsed; CR: complete remission.
In this study, we used both tests to assess the P-gp status in leukemic blasts. Peripheral blood or bone marrow samples from 50 acute leukemia patients were evaluated. Fourteen samples (28%) scored positive for the functional JC-1 test. A reversible rho123 efflux was seen in 18 samples (36%). In 2 patients, no rho123 transport was detected although they were clearly positive for the JC-1 test. These patients had blasts with a very low P-gp activity. However, even low P-gp activity can cause multidrug resistance. Six patients scored positive for the rho123 test, but were negative for the JC-1 assay. In one of these samples, the percentage of leukemic blasts with a reversible JC-1 efflux was just below the threshold. In the other samples, the percentage of leukemic blasts showing a reversible rho123 efflux was very low. The leukemic cells in these samples were negative for the immunological MRK16 assay which means that they did not express P-glycoprotein. Therefore, these results were considered to be false positive. The rho123 is probably transported by MRP1 in stead of P-gp [24].

We found a strong correlation between the functional JC-1 assay and the rho123 test. This is in contrast with the results of Legrand et al. [15]. They found no correlation between the rho123 efflux and the percentage of leukemic blasts showing a reversible shift in red fluorescence. This difference is probably due to the fact that we optimized our rho123 assay by using tumor specific monoclonal antibodies labeled with allophyco-cyanin to identify the leukemic blasts.

AML patients exhibited a significantly higher level of P-gp function compared to ALL patients. Eleven out of 24 AML patients (46%) scored positive for the functional JC-1 assay. Fourteen AML patients (58%) showed a reversible rho123 efflux. Although this is less than the
percentage of positive AML patients found by Legrand et al. (65%) [15], it is still comparable with previously published results [2,26,27].

We studied the association between the P-gp status and the clinical outcome in 30 children and 20 adults suffering from acute leukemia. In the pediatric group, there was no correlation between P-gp activity and clinical outcome. This is in accordance with the results of Kanerva et al. and Wuchter et al. but in conflict with the data of De Moerloose et al. [5,6,28]. This lack of correlation was probably due to small patient group and the heterogeneity of the samples. In addition, multidrug resistance can also be caused by the overexpression of other transport pumps such as the members of the multidrug resistance associated protein family (MRP 1–6) and the lung cancer related protein (LRP). Further investigations will be necessary to determine the prognostic value of these pumps in childhood acute leukemia.

We also analyzed the P-gp function in two samples taken at different time points from a child with AML. The first sample, taken at relapse, scored negative for the JC-1 assay. In a second sample, taken three weeks later, 20% of the leukemic cells showed a reversible JC-1 efflux. This is in accordance with previously published results showing that the treatment of leukemic patients with anthracyclines and vinca alkaloids may further increase the P-gp expression in previously negative patients since these drugs are substrates of the P-gp pump [29–31].

Despite the small patient group and the heterogeneity of the samples, we demonstrated a correlation between the JC-1 status and the clinical outcome in the adult acute leukemia patients. This is in accordance with the results of Lamy et al., Leith et al. and Legrand et al. [2,3,26].

In conclusion, we highly recommend the JC-1 assay to study the P-gp activity in acute leukemia because it is more specific and less labor intensive than the conventional functional flow cytometric test using rhodamine 123. In addition, our results show that the functional JC-1 assay can be used to identify adult patients with an increased risk for an adverse clinical outcome.

Acknowledgments

The authors thank Prof. Dr F. Offner, Dr A. Janssens, Dr M. Petrick and Dr H. Louagie for sending samples from adult acute leukemia patients. We thank D. Claeys and M. Petrick and Dr H. Louagie for sending samples from the Flemish Institute for the Promotion of Scientific Technological Research in Industry (I.W.T).

References


4. Discussion and future perspectives

A major cause of treatment failure in childhood ALL is pre-existent or acquired resistance to a variety of structurally and functionally unrelated chemotherapeutic compounds. Generally, MDR is associated with transmembrane protein-mediated efflux of cytotoxic compounds leading to a decreased intracellular drug accumulation and toxicity. Several MDR-related drug efflux pumps have been characterized, including P-gp, MRP1, BCRP and LRP. Current knowledge about the clinical relevance and prognostic significance of multidrug resistance-related proteins in childhood ALL is reviewed in Paper 1. However, the lack in standardized detection techniques, the differences in treatment protocols and the heterogeneity in patient groups (e.g. pooled data of ALL and AML, initial and relapse samples, adults and children), hamper the comparison of data.

Several authors reported a relationship between P-gp expression and/or function and clinical outcome in children with ALL.\textsuperscript{168-177} Other groups, however, contradicted these findings.\textsuperscript{178-181} We assessed the clinical relevance of P-gp expression and activity in childhood ALL in a prospective study (Paper 2). The P-gp expression was analyzed using an immunocytochemical assay and the P-gp activity was measured by a functional flow cytometric test using rho123. Neither P-gp expression nor P-gp activity were associated with a higher relapse rate or an increased risk of death. However, the combination of P-gp expression and activity revealed a statistically significant parameter predicting relapse in childhood ALL. Furthermore, patients with a positive result for both tests had a 4.5-fold higher risk of death compared to the negative group. Based on our findings, we recommend combining complementary techniques such as immunocytochemistry and flow cytometry in order to guarantee the accurate assessment of P-gp expression and activity. Moreover, pediatric ALL patients with increased risk for relapse can only be identified in case both tests are combined.

Information on the prognostic significance of other MDR proteins is scarce. LRP and BCRP might contribute to drug resistance in precursor B-cell lymphoblastic leukemia but these findings are not confirmed yet.\textsuperscript{183-187} MRP1 does not seem to play a major role in clinical resistance in pediatric ALL patients\textsuperscript{176,184,186} and of all newly identified ABC transporters (e.g. MRP2-MRP9, sP-gp and TAP), only MRP3 expression is associated with a significantly worse prognosis.\textsuperscript{188}
One of the reasons for the variability in published data is the lack in standardized, sensitive and specific detection techniques. Although this problem has been addressed by a number of international workshops, development of consensus recommendations has been difficult due to differences in assay sensitivity and specificity, the need to distinguish between normal and malignant cells, and the definition of a clinically relevant threshold.\textsuperscript{189-193} However, functional flow cytometric tests assessing modulator-induced changes in fluorophore retention and/or efflux are promoted because they allow, in contrast to immunological or molecular tests, the evaluation of protein activity (Table 3).

<table>
<thead>
<tr>
<th>MDR-related protein</th>
<th>Substrate</th>
<th>Inhibitor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P-gp</strong></td>
<td>Rhodamine 123</td>
<td>Verapamil</td>
</tr>
<tr>
<td></td>
<td>JC-1</td>
<td>Cyclosporin A</td>
</tr>
<tr>
<td></td>
<td>DiOC\textsubscript{2}</td>
<td>PSC833</td>
</tr>
<tr>
<td></td>
<td>Calcein AM</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Syto16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Daunorubicin</td>
<td></td>
</tr>
<tr>
<td><strong>MRP</strong></td>
<td>CFDA</td>
<td>Buthionine Sulfoximine</td>
</tr>
<tr>
<td></td>
<td>Calcein-AM</td>
<td>MK-571</td>
</tr>
<tr>
<td></td>
<td>Rhodamine 123</td>
<td>Probencid</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Genistein</td>
</tr>
<tr>
<td><strong>BCRP</strong></td>
<td>Mitoxantrone</td>
<td>Fumitremorgin C</td>
</tr>
<tr>
<td></td>
<td>Prazosin</td>
<td>GF120918</td>
</tr>
<tr>
<td><strong>LRP</strong></td>
<td>Daunorubicin</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

Legend: JC-1: 5,5',6,6'-tetrachloro-1,1',3,3'-tetraethylbenzimidazolcarbocyanine iodide; DiOC\textsubscript{2}: 3,3'-diethyloxacarbocyanine iodide; Calcein-AM: calcein acetoxy-methyl ester; CFDA: carboxy fluorescein diacetate

We compared a highly sensitive flow cytometric assay using the carbocyanine liquid crystal forming probe JC-1 with a conventional technique using rhodamine 123 (rho123) to assess P-gp activity in acute leukemia (Paper 3). Compared to rho123, JC-1 has several advantages: 1) since JC-1 does not cause background problems, up to three different tumor specific monoclonal antibodies can be used to identify leukemic blasts, 2) the JC-1 test does not require drug efflux monitoring which makes it much faster and less labor intensive than the conventional rho123 test, 3) JC-1 is not transported by MRP1 which increases the specificity of the test\textsuperscript{182}, and 4) the JC-1 assay can be used to detect low P-gp activity which is extremely important since even a weak efflux may lead to cellular drug resistance.\textsuperscript{182} Based on these findings, we decided to implement this novel sensitive assay in the lab.
In addition, a study evaluating the clinical importance of P-gp activity measured by JC-1 accumulation in childhood ALL is underway. From May 2001 till May 2004, peripheral blood or bone marrow samples, taken at diagnosis or relapse, from 30 children suffering from ALL have been analyzed using the functional JC-1 assay. Preliminary data indicate that there is no association between the P-gp activity measured by JC-1 accumulation and the clinical outcome of these children. However, continued accrual and longer follow-up is necessary before a definitive conclusion can be drawn.

Detailed information on the nature, function, substrate and inhibitor specificity, and clinical relevance of MDR-related proteins is needed before the potential of transporter-specific modulators can be explored. So far, phase 3 clinical trials with P-gp antagonists in adult AML patients have yielded conflicting results.\textsuperscript{194,195} This may be explained by the co-expression of other cellular drug resistance efflux pumps. The different transporters may act in concert and inhibition of one transporter may lead to compensation by other transport proteins. This functional redundancy between different MDR-related proteins represents a challenge in clinical studies evaluating correlations between drug sensitivity and expression of specific pumps. In our opinion, large and well controlled multicenter studies are needed to determine the prognostic significance of MDR-related transport proteins in childhood ALL. In addition, an international effort should be made to develop, optimize and standardize sensitive and specific detection techniques.

In conclusion, the JC-1 assay, evaluated and optimized within the framework of this thesis, is an asset for studies evaluating the clinical importance of P-gp activity. In addition, the combination of P-gp expression and activity predicts relapse in childhood ALL. However, large multicenter, quality controlled, prospective clinical studies are needed to confirm our results and elucidate the clinical relevance of other MDR-related proteins.
Chapter 3

Minimal Residual Disease
1. INTRODUCTION..........................................................................................................73
2. DETECTION OF DISSEMINATED NEUROBLASTOMA CELLS .................................74
   2.1. PAPER 4..............................................................................................................74
   2.2. PAPER 5..............................................................................................................92
   2.3. PAPER 6............................................................................................................104
   2.4. PAPER 7............................................................................................................126
3. CLINICAL IMPORTANCE OF NEUROBLASTOMA MICROMETASTASES ..........151
   3.1. STAGING AND RISK ASSESSMENT AT DIAGNOSIS ........................................151
   3.2. MONITORING DISEASE COURSE AND RESPONSE TO THERAPY .....................151
   3.3. SCREENING OF AUTOLOGOUS STEM CELL PREPARATIONS .............................152
4. DISCUSSION AND FUTURE PERSPECTIVES........................................................153
1. Introduction

At diagnosis, approximately 40% of the neuroblastoma patients present with high risk stage 4 disease and bone marrow involvement. These children have a poor clinical outcome despite intensive multimodal therapy including autologous stem cell transplantation. According to the INSS, conventional cytomorphological screening of bone marrow smears is still the only accepted method for the detection of disseminated neuroblastoma cells. However, Méhes and colleagues reported that a tumor cell infiltrate below 0.1% can be overlooked by conventional cytomorphology. Consequently, more sensitive and specific detection techniques are urgently needed in order to guarantee correct clinical staging and risk assessment at diagnosis. These methods may also be applied to monitor therapeutic response during treatment and detect relapse before overt metastasis occur, resulting in earlier clinical intervention. In addition, screening of autologous stem cell preparations is crucial since the re-infusion of contaminated stem cell products could lead to systemic recurrence.

Within the framework of this thesis, three new detection techniques were developed and evaluated. In collaboration with the SIOPEN Bone Marrow Subcommittee, an immunocytochemical assay, based on the detection of GD2 disialoganglioside, was optimized and standardized (Paper 4). Moreover, morphological and immunological evaluation criteria were formulated and a work flow was designed. In paper 5, two four-color flow cytometric assays using different combinations of CD9, CD81, CD56, CD45 and anti-GD2 disialoganglioside are described. Disseminated neuroblastoma cells were also detected by means of TH, ELAVL4 and GD2 synthetase real-time quantitative RT-PCR (QPCR) (Paper 6). Current knowledge about the detection techniques, used to study minimal residual disease, is reviewed in paper 7.
2. Detection of disseminated neuroblastoma cells

2.1. Paper 4

STANDARDIZATION OF THE IMMUNOCYTOCHEMICAL DETECTION OF NEUROBLASTOMA CELLS IN BONE MARROW


Standardization of the immunocytochemical detection of neuroblastoma cells in bone marrow


ABSTRACT

Standard cytomorphological examination of bone marrow (BM) aspirates appears not sensitive enough to detect single neuroblastoma cells. The SIOPEN Neuroblastoma Bone Marrow Committee developed a sensitive and reproducible anti-GD2 immunocytochemical assay and introduced morphological and immunocytological criteria for the interpretation of results. Fixed cytopsins were incubated with a commercially available anti-GD2 monoclonal antibody and an APAAP kit. Cells fulfilling all morphological and immunocytological criteria were called Criteria Positive Cells (CPC’s). Not Convincingly Interpretable Cells (NCIC’s) fulfilled some but not all criteria and Negative Cells (NC’s) displayed only exclusion criteria. The genetic profile of doubtful cells was checked by FISH. Ideally, 3x10^6 cells were analyzed to reach a 95% probability of detecting one tumor cell in 1x10^6 mononuclear cells. Four quality control (QC) rounds were organized to validate the method. A total of 111 QC samples were analyzed. Two main improvements were achieved: in discordant cases, the range between lowest and highest reported result was reduced by half, and discordant results were only found in samples with less than 10 CPC’s per 1x10^6. This paper describes the first internationally standardized protocol to detect and quantify rare neuroblastoma cells by immunocytochemistry. This method is an indispensable tool for multicenter studies evaluating the clinical significance of minimal residual disease in neuroblastoma.
INTRODUCTION

Neuroblastoma (NB), a tumor originating from the sympathetic nervous tissue, is the most common extra-cranial solid tumor in children with a yearly incidence of 7 to 10 per million. The tumor consists of sympathetic neuronal elements of variable immaturity and shows a diverse clinical behavior. Approximately 40% of the NB patients suffer from high risk stage 4 disease with bone marrow (BM) involvement. These children have a poor clinical outcome despite intensive multimodal therapy.

The demonstration of disseminated tumor cells in BM is important for clinical staging and risk assessment at diagnosis and for monitoring therapeutic response during treatment. In addition, screening of autologous stem cell preparations is crucial since the re-infusion of contaminated stem cell products could lead to systemic recurrence.

According to the International Neuroblastoma Staging System (INSS), conventional cytology of BM smears is still the only accepted technique for the detection of disseminated NB cells. However, the sensitivity of this approach is limited since a tumor cell number below 0.1% is virtually not detectable by conventional cytomorphology. Therefore, the development of more sensitive and specific detection methods is indispensable.

During the last decades, several assays based on immunocytochemistry, automatic immunofluorescence plus FISH (AIPF), reverse transcription polymerase chain reaction (RT-PCR) or flow cytometry were evaluated. However, the reliability of tumor cell detection and quantification by these methods remains controversial.

A Neuroblastoma Bone Marrow Committee (NBMC) was established by the European Neuroblastoma Study Group to evaluate and standardize procedures for the detection of minimal residual disease (MRD) in NB patients. In connection with the evaluation of a new high risk protocol by the SIOP European Neuroblastoma (SIOPEN) group, the NBMC developed, optimized and standardized an immunocytochemical assay based on the detection of the neuroblastoma specific GD2 disialoganglioside. In addition, morphological and immunocytological criteria for the interpretation of results were introduced and standardized. Four multicenter quality control rounds were organized among nine European research groups to evaluate the technique and assess the inter-observer concordance. The latter improved markedly after the adoption of the standardized protocol.

This paper describes a standardized immunocytochemical staining method and minimal morphological and immunocytological criteria for the evaluation of stained BM samples. The application of this protocol will lead to a more reproducible and reliable assessment of MRD in NB. We believe that a standardized method is needed to generate comparable results in multicenter studies evaluating the clinical significance of MRD.
MATERIALS AND METHODS

Sample collection
After informed consent, bilateral BM aspirates from the iliac crest were performed following previously published guidelines.\textsuperscript{6,21} The second aspiration from each puncture site was used for immunocytology and FISH. Bilateral BM samples were not pooled and were transferred to the laboratory at room temperature as fast as possible.

Control samples
Slides containing cells from a NB cell line (e.g. IMR32) were included in every experiment as positive controls. For negative control, the primary anti-GD2 antibody was replaced with an antibody of the same IgG2a isotype (Dako Corporation, Glostrup, Denmark). This allowed us to evaluate the background staining caused by the interaction of the anti-GD2 antibody with Fc-receptor bearing leucocytes.

Isolation, processing and storage of mononuclear cells
Mononuclear BM cells were isolated by density gradient centrifugation using Lymphoprep (Nycomed, Oslo, Norway) following the instructions of the manufacturer. Aspirates from different sites were processed separately. After separation, the mononuclear cells were collected from the interphase layer and washed twice in phosphate buffered saline (PBS) (Gibco, Paisley, UK).
Large diameter cytospins (17 mm) containing approximately 5x10\textsuperscript{5} mononuclear cells were prepared. These slides should not be overcrowded and the mononuclear cells must lie well separated from one another. This can be achieved by centrifuging no more than 7x10\textsuperscript{5} cells down on precoated (e.g. poly-L-lysine) glass slides in a Hettich centrifuge (Hettich Zentrifugen, Tuttlingen, Germany).
The slides were air-dried overnight and stored in air-tight plastic boxes or wrapped in aluminum foil at \textdegreeC -24 \textdegreeC until immunocytology was performed. Before staining, the slides were thawed in closed boxes to avoid formation of condensation water because this could destroy the morphology of the cells.

Standardized immunocytological staining protocol
The NBMC decided to standardize the fixation, the immunocytochemical staining procedure and the evaluation of immunocytological results in order to improve the sensitivity, specificity and reproducibility of MRD detection in NB.
Fixation:
The cytospins must be fixed in 4 % buffered paraformaldehyde for 10 minutes. Commercially available formaldehyde can also be used provided that it is free from methanol. In order to avoid artificial tumor cell contamination, slides must be incubated individually. After fixation, the cytospins are washed three times with PBS to remove remainders of the fixative.

Immunocytochemical staining:
The staining procedure comprises 10 subsequent steps. All incubations are performed in a humidifier at room temperature.

1. Incubation of cytospins for 30 minutes with 30 µl of an unlabeled monoclonal mouse anti-human GD2 disialoganglioside antibody (clone 14.G2a, BD Biosystems, Erembodegem, Belgium) diluted 1/100 in 1 % BSA in PBS (Bovine Serum Albumin, Gibco, Paisley, UK).
2. Washing twice with PBS for 5 min.
3. Incubation for 30 minutes with 30 µl of an unlabeled rabbit anti-mouse antibody (Dako Corporation, Glostrup, Denmark), diluted 1/20 in 1 % BSA in PBS.
4. Repetition of step two.
5. Incubation with 30 µl of the APAAP complex (Dako Corporation, Glostrup, Denmark), diluted 1/20 in 1 % BSA in PBS.
6. Repetition of step two.
7. Incubation with Dako Fuchsin+ ™ Substrate Chromogene System (Dako Corporation, Glostrup, Denmark), prepared as indicated by the manufacturer, for no longer than 10 minutes.
8. Washing in running tap water for at least 5 minutes.
9. Counterstaining with hematoxylin (Sigma, St. Louis, USA) until an appropriate blue nuclear stain is obtained.
10. Mounting in aqueous mounting medium (e.g. Glycergel mounting medium, Dako Corporation, Glostrup, Denmark) with coverslip.

Evaluation of immunocytochemically stained samples:
The characteristics of GD2-positive disseminated NB cells and false positive hematopoietic cells were examined in detail. The observations lead to the following morphological and immunocytological guidelines for the identification of positive cells.
**Morphological criteria:**
Cells with a round nucleus, often, but not always larger than that of small lymphocytes, displaying a granular chromatin structure and a scarce amount of cytoplasm are considered positive. Cells showing a low nuclear/cytoplasmic ratio or typical morphological features of hematopoietic cells are considered negative. Cells found outside or on the boarder of the cytospin field are excluded.

**Immunocytological criteria:**
Cells must display a strong, deep red staining localized to the entire cell membrane and cytoplasm. A weak staining and a staining restricted to a subcellular compartment or covering the nucleus is considered negative. In addition, cells surrounded by positively stained amorphous material are excluded.

Based on the morphological and immunocytological criteria, cells are classified into three groups:
1) CPC’s (Criteria Positive Cells): cells fulfilling all morphological and immunocytological inclusion criteria (Figure 1 A-D).
2) NCIC’s (Not Convincingly Interpretable Cells): cells fulfilling some but not all inclusion criteria (Figure 1 E).
3) NC (Negative Cells): cells displaying only exclusion criteria (Figure 1 F-G).

Single cells as well as cells being part of a Homer Wright rosette or cell clump are evaluated, classified and counted. Clumps or rosettes consisting of too much cells to evaluate or count, are reported separately. In addition, the estimated number of evaluated mononuclear cells must be reported.

When evaluating or reporting immunocytochemical results, the work flow depicted in Figure 2 should be followed. No further review is needed when no GD2 positive cells or more than 10 CPC’s are present. For samples with 1 to 10 CPC’s or samples containing NCIC’s, central review by the members of the NBMC is obligatory. When no consensus is reached, the genetic profile of the doubtful cells should be checked by FISH to find out whether it corresponds to the cytogenetic aberrations found in the primary tumor. The sequential immunocytological staining and molecular cytogenetic characterization can be done using an automated scanning and relocation system (e.g. Metafer4/RCDetect, MetaSystems, Altluessheim, Germany). GD2 positive cells showing genetic aberrations (e.g. gain of whole chromosomes or 17q and MYCN amplification) are called FPC’s (FISH Positive Cells) (Figure 3).
**Figure 1:** Immunocytochemical analysis of BM slides from NB patients according to the standardized protocol. A: CPC’s (Criteria Positive Cells) forming a clump. NB cells in clumps do not always display a round nucleus because they adjust their form to the clump. The entrapped erythroblast (arrow) is smaller than the NB cells and displays a different chromatin structure. B: The 2 CPC’s mold in a different way compared to the surrounding hematopoietic cells. C: Two CPC’s showing the typical nuclear size, chromatin structure and nuclear/cytoplasmic ratio which clearly differs from hematopoietic cells. The myeloid cells (arrow) are passively stained due to the shedding of the GD2 antigen by the NB cells. D: If the morphological and immunological criteria are fulfilled, even a single cell can be identified as CPC. E: The staining intensity of this cell is comparable to that of a CPC but the nuclear shape, size, chromatin structure and the nuclear cytoplasmic ratio do not fulfill the criteria. Therefore the cell is classified as NCIC (Not Convincingly Interpretable Cell). F: Neither the nuclear/cytoplasmic ratio nor the vesicular structure of the cytoplasm fulfills the criteria. The cell has the cytological features of a histiocyte. G: The cell on the left side has the same size and displays the same staining intensity as the CPC on the right. However, the nuclear shape and the nuclear/cytoplasmic ratio do not fulfill the criteria. Moreover, the cell contains GD2 positive material in the cytoplasm (arrow). This is a typical feature of a macrophage. H: Strongly positive material (right) without a visible nucleus is not reported. CPC on the left.
**Figure 2:** Work Flow (CPC’s: Criteria Positive Cells; NCIC’s: Not Convincingly Interpretable Cells; MNC’s: mononuclear cells; AIPF: Automatic Immunofluorescence plus FISH; FPC’s: FISH Positive Cells)

- Negative sample > 10 CPC’s per 1x10^6 MNC’s
  - No further review
- Negative sample 1 to 10 single CPC’s per 1x10^6 MNC’s
  - Central review
  - FPC’s (AIPF)
- Only NCIC’s present

**Figure 3:** Automatic Immunofluorescence plus FISH (AIPF) allows the sequential immunological staining and molecular cytogenetic characterization of disseminated NB cells. A: GD2 positive cells (green fluorescence). The nucleus is stained with DAPI (blue fluorescence). B: The genetic make up of the GD2 positive cells is visualized by FISH. Only one GD2 positive cell displays MYCN amplification (green fluorescence). GD2 positive cells displaying the same genetic aberrations as the primary tumor are called FISH Positive Cells (FPC’s).
RESULTS

Sensitivity
The sensitivity of the immunocytochemical assay is not limited by the technique itself. On the contrary, the number of analyzed cells defines the sensitivity. When enough cells are analyzed, a high sensitivity can be reached. The Poisson Distribution \( f(X) = \mu^X e^{-\mu}/X! \) can be used to calculate the statistics of tumor cell detection.\(^{11}\) The parameter \( \mu = n.p \) (where \( n \) is the total number of cells analyzed and \( p \) is the true frequency of tumor cells) represents the total number of tumor cells present. The variable \( X \) denotes the number of tumor cells actually detected. When \( X = 0 \), \( f(X) \) is the possibility of missing a tumor cell. When only one tumor cell is present in the midst of 999999 normal cells, the probability of missing the tumor cell after counting \( 3 \times 10^6 \) cells is less than 5%. This means at least \( 3 \times 10^6 \) cells should be analyzed to reach a 95% chance of detecting one tumor cell in \( 1 \times 10^6 \) normal mononuclear cells. Therefore, six cytospins, each containing \( 5 \times 10^5 \) cells, should be analyzed to secure the analysis of \( 3 \times 10^6 \) mononuclear cells.

Quality Control
From 2001 to 2003, 4 multicenter quality control (QC) rounds were organized among the nine members of the NBMC to develop and validate the staining protocol, the morphological and immunocytological criteria and the work flow. A total of 111 QC samples were analyzed. Every research group sent preferably three slides from at least two different BM samples to every other member. The participants fixed and stained the cytospins independently of each other. They evaluated the samples unaware of any clinical information. Individual screening results were disclosed and the level of inter-observer concordance was assessed during subsequent quality control meetings. In addition, samples with discordant results were reviewed by all participants resulting in an optimized staining protocol and refined morphological and immunocytological criteria. The results of quality control round 1 and 4 are shown in Table 1. Only samples analyzed by at least four participants were included in the study. Participants 1 to 8 used the immunocytochemical staining assay to evaluate the quality control samples. Participant 9 detected residual NB cells using automatic immunofluorescence plus FISH.

Quality control round 1 was organized before the immunocytochemical staining protocol was standardized and the morphological and immunocytological criteria were formulated. A total of 33 QC samples were analyzed. Two samples were excluded because they were analyzed by only three participants. Considerable differences were found both in the number of positively scored samples and in the number of GD2 positive cells per individual sample. Six samples were scored positive by all participating centers. In five samples, no positive cells
were found. Discordant results were found in 20 samples (65%). In these samples, the average difference between the highest and the lowest reported number of GD2 positive cells was 19.

Quality control round 4 was organized after the standardization of the staining method, the formulation of the criteria and the design of the workflow. Thirteen samples were fixed, stained and evaluated by each participant in a blinded way and, in accordance with the workflow, samples containing less than 10 CPC’s or samples with NCIC’s were reviewed during a quality control meeting. Since we noticed a remarkable improvement in the sensitivity and specificity of the method and in the reproducibility of the results after analyzing 13 samples, QC round 4 was terminated at that point. After central review, the results were concordant in 10 samples. Only in 3 samples (23 %) discordant results were found. In these samples, the average difference between the highest and the lowest reported number of CPC’s was 9. Discordant results were only found in samples with less than 10 CPC’s per 1x10^6.

The standardization of the assay led to a significant decrease in the number of discordant results (χ^2 = 4.91, p = 0.027, dF = 1). The range between the highest and the lowest reported number of positive cells decreased from 19 to 9.

The immunocytochemical results (participants 1 to 8) were also compared to those obtained with AIPF (participant 9). In quality control round 1, 15 samples scored positive for the immunocytochemical assay whereas no neuroblastoma cells were detected by AIPF. The discrepancies in four of these samples were probably due to sample variability since only one or two IC positive cells were found. All other samples (11) were most likely false positive. In QC round 4 only one discordant result was found (8%). Participant 7 reported 8 CPC’s whereas no FPC’s were detected by participant 9. These results prove that the standardization of the staining and evaluation procedures reduced the number of false positive results dramatically.
**Table 1:** Results of quality control rounds 1 and 4. Participants 1 to 8 used an immunocytochemical staining assay to evaluate the quality control samples. Participant 9 detected residual neuroblastoma cells using an automatic immunofluorescence plus FISH device (AIPF). The number of GD2 positive cells (quality control round 1) or the number of CPC's (quality control round 4) is reported. (NE: not evaluable; ND: not done; C: 1 or 2 positive clusters; >100: more than 100 positive cells or more than 2 clusters)

<table>
<thead>
<tr>
<th>Cytospin Label</th>
<th>Lab 1</th>
<th>Lab 2</th>
<th>Lab 3</th>
<th>Lab 4</th>
<th>Lab 5</th>
<th>Lab 6</th>
<th>Lab 7</th>
<th>Lab 8</th>
<th>Lab 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>QC1 A1</td>
<td>NE</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>NE</td>
<td>NE</td>
<td>0</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 A2</td>
<td>NE</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>NE</td>
<td>0</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 A3</td>
<td>NE</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>NE</td>
<td>0</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 A4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>1</td>
<td>0</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 B1</td>
<td>6</td>
<td>6</td>
<td>9</td>
<td>5</td>
<td>&gt;100</td>
<td>53</td>
<td>42</td>
<td>ND</td>
<td>10</td>
</tr>
<tr>
<td>QC1 B2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>3</td>
<td>15</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 B3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>ND</td>
<td>3</td>
<td>1</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 C1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>NE</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>ND</td>
<td>&gt;100</td>
</tr>
<tr>
<td>QC1 C2</td>
<td>50</td>
<td>&gt;100</td>
<td>NE</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>ND</td>
<td>&gt;100</td>
</tr>
<tr>
<td>QC1 C3</td>
<td>NE</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 G1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>NE</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>ND</td>
<td>&gt;100</td>
</tr>
<tr>
<td>QC1 G2</td>
<td>C</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>NE</td>
<td>0</td>
<td>15</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 G3</td>
<td>NE</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NE</td>
<td>0</td>
<td>18</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 G4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>NE</td>
<td>0</td>
<td>21</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 I1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>NE</td>
<td>NE</td>
<td>&gt;100</td>
<td>NE</td>
<td>&gt;100</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>QC1 I2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>NE</td>
<td>0</td>
<td>NE</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>QC1 I3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>ND</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>QC1 I4</td>
<td>11</td>
<td>C</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>32</td>
<td>ND</td>
<td>7</td>
</tr>
<tr>
<td>QC1 I5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 I6</td>
<td>40</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>ND</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>ND</td>
<td>&gt;100</td>
</tr>
<tr>
<td>QC1 I7</td>
<td>0</td>
<td>2</td>
<td>ND</td>
<td>0</td>
<td>ND</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 I8</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>2</td>
<td>ND</td>
<td>0</td>
<td>8</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 N1</td>
<td>0</td>
<td>&gt;100</td>
<td>C</td>
<td>C</td>
<td>ND</td>
<td>23</td>
<td>92</td>
<td>ND</td>
<td>80</td>
</tr>
<tr>
<td>QC1 N2</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>2</td>
<td>ND</td>
<td>0</td>
<td>8</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 N3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>ND</td>
<td>7</td>
<td>14</td>
<td>ND</td>
<td>7</td>
</tr>
<tr>
<td>QC1 S1</td>
<td>0</td>
<td>ND</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>NE</td>
<td>16</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 S2</td>
<td>0</td>
<td>ND</td>
<td>0</td>
<td>0</td>
<td>NE</td>
<td>18</td>
<td>ND</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>QC1 S3</td>
<td>0</td>
<td>ND</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>QC1 S4</td>
<td>0</td>
<td>ND</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>QC1 S5</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>ND</td>
<td>1</td>
<td>22</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>QC1 S6</td>
<td>C</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>ND</td>
<td>0</td>
<td>27</td>
<td>ND</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cytospin Label</th>
<th>Lab 1</th>
<th>Lab 2</th>
<th>Lab 3</th>
<th>Lab 4</th>
<th>Lab 5</th>
<th>Lab 6</th>
<th>Lab 7</th>
<th>Lab 8</th>
<th>Lab 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>QC4 A1</td>
<td>0</td>
<td>0</td>
<td>NE</td>
<td>NE</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>QC4 A2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>QC4 B1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>QC4 B2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>QC4 C1</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>ND</td>
<td>&gt;100</td>
</tr>
<tr>
<td>QC4 C2</td>
<td>0</td>
<td>C</td>
<td>ND</td>
<td>ND</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>ND</td>
<td>4</td>
</tr>
<tr>
<td>QC4 F1</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>QC4 F2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>QC4 G1</td>
<td>48</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>&gt;100</td>
</tr>
<tr>
<td>QC4 N1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>QC4 N2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>QC4 S1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>QC4 S2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

84
DISCUSSION

The detection of occult NB cells in BM has important therapeutic and prognostic implications since BM disease is associated with an unfavorable outcome for most children.\textsuperscript{22,23} On the other hand, children with stage 4S disease have a good prognosis although they may present with BM metastases. Cytomorphological screening of BM smears is still the only accepted method for the detection of disseminated NB cells. However, Méhes et al. reported that a tumor cell infiltrate below 0.1% can be overlooked by conventional cytomorphology because the unspecific morphological appearance of NB cells limits the sensitivity of this method.\textsuperscript{7} During recent years, numerous alternative approaches using immunological and molecular biological techniques were developed to improve the detection of residual NB cells in BM.\textsuperscript{9-20} However, the sensitivity and specificity of these assays vary markedly and hamper studies evaluating the clinical significance of MRD. Therefore, the standardization of detection techniques is urgently needed.

In connection with a phase III study organized by the SIOP European Neuroblastoma (SIOPEN) group, the NBMC developed, optimized and standardized an immunocytochemical assay based on the detection of the NB specific GD2 disialoganglioside. This antigen is highly and consistently expressed in neuroectodermal tumors and is not found in normal BM or peripheral blood cells.\textsuperscript{11,24,25} Since all NB are believed to express GD2, the possibility of a false negative result can practically be excluded. However, the uneven distribution of NB cells in the body may hamper the detection of disseminated tumor cells. To avoid false negative results caused by sampling error, a sufficient number of cells must be analyzed. The Poisson Distribution $f(X) = \mu^X \cdot e^{-\mu}/X!$ can be used to calculate the statistics of tumor cell detection.\textsuperscript{11} At least $3 \times 10^6$ cells must be analyzed to reach a 95\% probability of detecting one tumor cell in $1 \times 10^6$ normal mononuclear cells.

Immunocytochemical results can also be obscured by false positive events.\textsuperscript{26-28} These can be caused by the active or passive take up of the tumor cell derived GD2 ganglioside by hematopoietic cells. Furthermore, a very small subset of mature plasma cells producing antibodies against alkaline phophatase can be false positive since they react directly with the enzyme.\textsuperscript{27} Finally, the illegitimate expression of the targeted antigen, the cross-reactivity of the applied monoclonal antibody and interactions between antibodies and Fc-bearing leukocytes can give rise to false positive results.

As long as the members of the NBMC stained their slides using different immunocytochemical staining methods and analyzed their results according to individual morphological criteria, considerable discrepancies were observed. The evaluation of stained samples by the whole group using a multiheaded microscope, clearly demonstrated the
urgent need for developing a standardized immunocytochemical staining protocol and introducing morphological and immunocytological criteria. Consequently, the NBMC agreed upon one staining method and formulated morphological and immunocytological criteria for the interpretation of the results. Only cells with a round nucleus, often, but not always larger than that of small lymphocytes, a granular chromatin and a limited amount of cytoplasm are considered positive. In addition, a strong, deep red staining localized to the entire cell membrane and cytoplasm must be present. To our knowledge, this is the first time that a standardized protocol including morphological and immunological criteria for the detection of NB cells in BM has been designed.

However, when applying these criteria the NBMC discovered that a small proportion of immunocytochemically stained cells could not unequivocally be classified as positive (i.e. NB cells) because they did not fulfill all postulated morphological and immunocytological criteria. Therefore, it was decided to categorize all immunocytochemically stained cells into three groups: Criteria Positive Cells (CPC’s) fulfilling all postulated criteria, Not Convincingly Interpretable cells (NCIC’s) displaying some but not all inclusion criteria and Negative Cells (NC’s) which, in spite of their staining, were identified as non-malignant hematopoietic cells.

Borgen et al. published a similar approach for the analysis of circulating carcinoma cells by applying the anti-cytokeratin antibodies AE1/AE3 and an alkaline phosphatase-based detection method on cytospins prepared from mononuclear BM cells.\(^{29}\) They also categorized immunologically stained cells into three groups which they called tumor cells, probable tumor cells (?), and hematopoietic cells. In order to discriminate between these groups they presented a catalogue of pictures illustrating a large number of morphological and immunological variants of these categories. However, regarding the detection of NB cells using an anti-GD2 antibody, the members of the NBMC do not believe that it is feasible to cover all possible variants of immunocytochemically stained BM cells by means of illustrations. Therefore, they decided to introduce a workflow including two additional analytical steps (Figure 2). Firstly, samples with 1 to 10 CPC’s or samples containing NCIC’s are simultaneously reviewed by the members of the NBMC. Secondly, if no consensus is reached, the genetic profile of the doubtful cells is checked by automatic immunofluorescence plus FISH to disclose the identity of these cells. If the genetic aberrations in the doubtful cells correspond to those found in the primary tumor, the cells are called FISH positive cells (FPC’s). Finally, the morphological and immunocytological features of these FPC’s are carefully studied in order to refine the standardized evaluation criteria.

The standardized staining protocol, the morphological and immunocytological criteria and the work flow were evaluated during four multicenter quality control rounds, organized among the nine members of the NBMC. A total of 111 QC samples were analyzed. The concordance between the different observers, with regard to the staining and the evaluation of the
immunocytochemical results was assessed. After standardization, a significant decrease in the number of discordant results was reported. In addition, the range between the highest and the lowest reported result was reduced by half and discordant results were only found in samples with less than 10 CPC’s per 1x10^6 mononuclear cells.

Immunocytology has been used in the clinical practice of hematology and oncology for many years and has many advantages compared with flow cytometry or RT-PCR. In contrast to the latter, immunocytology allows the reliable quantification of tumor cells. This is important when the number of disseminated tumor cells appears to be prognostically important and not purely the presence or the absence of disease. The immunocytological technique is cost-effective and simple and since no expensive equipment is needed, immunocytological stainings can be performed in virtually every routine laboratory around the world.

This paper describes the first international standardization of an immunocytochemical staining and evaluation method developed to detect and quantify small numbers of neuroblastoma cells in BM. The results of our quality control rounds show that the standardization of the staining method, the formulation of morphological and immunocytological criteria and the design of the work flow resulted in a higher reproducibility, sensitivity and specificity. Methodological standardization is indispensable and must be agreed upon before multicenter studies, designed to assess the clinical importance of minimal residual disease, can be initiated.

ACKNOWLEDGMENTS

This work was supported by the SIOPEN-R-NET project (EC grant QLRI-CT-2002-01768). A part of this work was funded by a Ph.D. grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen).
REFERENCES


2.2. Paper 5

DETECTION OF RESIDUAL NEUROBLASTOMA CELLS IN BONE MARROW: COMPARISON OF FLOW CYTOMETRY WITH IMMUNOCYTOCHEMISTRY

Swerts K., De Moerloose B., Dhooge C., Brichard B., Benoit Y., Laureys G. and Philippé J.

Cytometry Part B (Clinical Cytometry) 2004; 61B: 9-19
Detection of Residual Neuroblastoma Cells in Bone Marrow: Comparison of Flow Cytometry With Immunocytochemistry

Katrien Swerts,1* Barbara De Moerloose,1 Catharina Dhooge,1 Bénédicte Brichard,2 Yves Benoit,1 Geneviève Laureys,1 and Jan Philippe3
1Department of Pediatric Hematology and Oncology, Ghent University Hospital, Ghent, Belgium
2Department of Pediatric Hematology and Oncology, Cliniques Universitaires Saint-Luc, Brussels, Belgium
3Department of Clinical Chemistry, Microbiology and Immunology, Ghent University Hospital, Ghent, Belgium

Background: Because the cytomorphologic examination of bone marrow (BM) aspirates appears not sensitive enough to detect residual neuroblastoma cells, two four-color flow cytometric assays using different combinations of CD9, CD81, CD56, CD45, and anti-GD2 were evaluated.

Methods: The sensitivity of the flow cytometric assays was assessed by spiking experiments in normal peripheral blood samples. Twenty-eight BM samples, 12 biopsies, and 3 peripheral blood stem cell (PBSC) preparations from 22 patients with neuroblastoma were analyzed. The results were compared with those of an anti-GD2 immunocytochemical reference assay.

Results: Flow cytometric and immunocytochemical analyses showed residual neuroblastoma cells in four BM samples. One PBSC preparation and 20 BM samples were negative for both assays. Four BM and two PBSC samples scored positive for the immunocytochemical assay but were negative for the flow cytometric tests. This was due to the limited number of cells that were flow cytometrically analyzed. A strong correlation between the flow cytometric and immunocytochemical tests was found ($\chi^2 = 6.4, P = 0.011$).

Conclusions: When an equal amount of cells is analyzed, the sensitivity of the flow cytometric assays is to be about 10 times lower than that of the immunocytochemical test. However, the flow cytometric assays can be used to screen for residual neuroblastoma cells in case of a GD2-negative primary tumor. Therefore we recommend flow cytometry for the detection of residual neuroblastoma cells. © 2004 Wiley-Liss, Inc.

Key terms: neuroblastoma; minimal residual disease; flow cytometry; immunocytochemistry

Neuroblastoma (NB), a tumor originating from the sympathetic nervous tissue, is the most common extracranial solid tumor in children. It shows a wide range of biologic, genetic, and morphologic characteristics and exhibits a diverse clinical behavior. Some children present with a localized tumor associated with a good prognosis, whereas others have a highly metastatic, aggressive tumor with an unfavorable outcome (1).

Approximately 40% of patients with NB have high-risk stage 4 disease with bone marrow (BM) involvement. These children have a poor clinical outcome, with a 5-year survival rate of approximately 30% (2). The presence of NB cells in the BM during therapy can predict relapse or clinical outcome (3–6). The detection of BM involvement is thus critical for accurate staging and risk assessment.

Myeloablative chemotherapy in combination with autologous BM or peripheral blood stem cell (PBSC) transplantation is commonly used to treat children with stage 4 disease (7). Residual NB cells can contaminate the stem cell product (8). This can cause the recurrence of the disease due to the reinfection of neuroblastoma cells (9).

Grant sponsor: Flemish Institute for the Promotion of Scientific Technological Research in Industry.
*Correspondence to: Katrien Swerts, Laboratory of Hematology (1P8 West), Ghent University Hospital, De Pintelaan 185, B-9000 Ghent, Belgium.
E-mail: katrien.swerts@ugent.be
Received 13 November 2003; Accepted 5 May 2004
Published online 26 July 2004 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/cyto.b.20019
Therefore, stem cell preparations must be screened for the presence of residual neuroblastoma cells (10).

The cytomorphologic examination of BM aspirates is not sensitive enough to detect residual NB cells (11). More sensitive assays, based on the use of immunocytochemistry, reverse transcription–polymerase chain reaction (RT-PCR) or flow cytometry, were developed to screen for NB cells (12–19).

Komada et al. evaluated a triple-color flow cytometric assay using CD9 in combination with CD56 and CD45 (12). This test was optimized by Nagai et al. (14) who combined CD81, CD56, and CD45. Warzynski et al. tested several combinations (18). They used CD45, CD56, a monoclonal antibody against neuron-specific enolase (NSE) and anti-GD2. GD2 disialoganglioside is highly expressed in almost all NBs and the GD2 density in NB cells is high (5 to 10^6 molecules/cell) (20–23). Warzynski et al. evaluated different double- and triple-color assays and analyzed anti-GD2, CD81, CD45, and CD56 expressions of the NB cell line SK-N-DZ by using a four-color flow cytometric assay that they currently use to evaluate their clinical samples (18).

In this study, we developed a new four-color flow cytometric assay combining CD9, CD81, CD56, and CD45. In addition, we evaluated CD81, CD56, and CD45 in combination with an anti-GD2 monoclonal antibody.

CD9 reacts with a 24-kDa type III transmembrane protein that is expressed on platelets, pre-B cells, activated and differentiating B cells, activated T cells, eosinophils, monocytes, endothelial and epithelial cells, brain, peripheral nerves, and muscular cells. Various in vitro studies have shown consistent expression of the CD9 antigen on NB cells (24,25).

The expression of CD81 in neuroblastoma cell lines has also been described (26). CD81 identifies a 26-kDa type III transmembrane protein (TAPA-1) that is involved in cell growth and signal transduction. It is broadly expressed on cells of hematopoietic origin. In addition, it is present on endothelial and epithelial cells.

CD56 reacts with the 175/220-kDa glycosylated antigen isoforms present on a subpopulation of peripheral blood, large, granular lymphocytes that demonstrate natural killer activity. It is also present on a subset of CD4+ and CD8+ T cells in peripheral blood, neural-derived cells and tumors, large granular lymphocyte and myeloid leukemias, small cell lung carcinomas, and myelomas. The expression of neural cell adhesion molecule in human neuroblastoma cell lines has been reported by Lipinski et al. (27).

GD2 disialoganglioside is expressed by neuroectodermally derived tumors such as melanoma, neuroblastoma, small cell carcinoma, and glioma. GD2 expression is not found in normal BM or peripheral blood samples (28).

CD45 reacts with the isoforms of the leukocyte common antigen present on all human white blood cells, including lymphocytes, monocytes, granulocytes, eosinophils, and thymocytes. It is not expressed in NB cell lines (29).

We analyzed 12 tumor samples from patients with NB and used the four-color flow cytometric assays to screen for residual NB cells in 28 BM samples and three PBSC preparations. The sensitivity of the flow cytometric assays was evaluated with spiking experiments. We assessed the specificity of our four-color flow cytometric tests by analyzing samples from a ganglioneuroma, a Wilms tumor, a lung sequestrum, a rhabdomyosarcoma, and a desmoplastic infantile gangliogioma (DIG). Wilms tumors and rhabdomyosarcomas, like NBs, Ewing sarcomas, non-Hodgkin lymphomas, retinoblastomas, medulloblastomas, some small cell carcinomas, and primitive neuroectodermal tumors, belong to the family of ‘small blue round cell’ neoplasms. Because the members of this family can often be morphologically confused with NB, the two four-color flow cytometric assays could be used to differentiate between NB and the other neoplasms.

The results of the flow cytometric assays were compared with those of an anti-GD2 immunocytochemical test, which was considered the reference method.

**MATERIALS AND METHODS**

**Patients and Sample Preparation**

Between May 2001 and August 2003, 12 biopsies, 28 BM samples, and 3 PBSC preparations from 22 patients with NB, taken at diagnosis or during treatment, were examined. In addition, samples of a ganglioneuroma, a Wilms tumor, a lung sequestrum, a rhabdomyosarcoma, and a DIG were analyzed. Patient characteristics are summarized in Table 1. Bilateral BM samples were considered as two samples because they were analyzed separately.

Clinical staging was done according to the International Neuroblastoma Staging System (1). The ethical committee approved the study, and informed consent was obtained from patients and/or their parents.

From fresh tumor samples, single-cell suspensions were prepared by using collagenase (1,000 U/ml in RPMI medium) for 1 h at 37°C. Subsequently, cells were washed and resuspended, and cell clumps were removed by filtering through a 70-μm nylon membrane.

Immediately after collection, mononuclear cells from BM samples and PBSC preparations were isolated by density gradient centrifugation on Ficoll-Hypaque.

All cells were kept at 4°C and analyzed within 24 h after collection.

**Control Samples**

The NB cell line CLB-GA (kindly provided by F. Speleman, Medical Genetics, Ghent University Hospital, Ghent, Belgium) was included in the experiments as a positive control. BM samples from 10 adults without malignant disease were used as negative controls and were analyzed with the flow cytometric and immunocytochemical tests.

**Monoclonal Antibodies**

NB cells were detected by a four-color flow cytometric assay using five different monoclonal antibodies (CD9, CD81, CD56, anti-GD2, and CD45). We evaluated CD9 (clone P1/33/2) labeled with fluorescein isothiocyanate (FITC), CD81 (clone JS-81) labeled with phycoerythrin (PE), CD45 (clone 2D1) labeled with peridinin chloro-
phyll protein (PerCP), CD56 (clone N901) labeled with allophycocyanin (APC), and an unlabeled monoclonal antibody that interacts with GD2 disialoganglioside (clone 14G2a).

CD9 was purchased from Dako Corporation (Glostrup, Denmark); CD81, CD45, and anti-GD2 were obtained from BD Biosystems (Erembodegem, Belgium); and CD56 was purchased from Beckman Coulter (Fullerton, CA, USA).

Flow Cytometric Assay

Mononuclear cells (5 × 10^5 cells) isolated from BM samples, PBSC preparations, or fresh tumor samples were resuspended in 100 μl RPMI medium and incubated for 30 min with an unlabeled isotypic control antibody immunoglobulin G2a as a control for the GD2 assay (5 μl; Dako Corporation). In parallel, cells were incubated with 5 μl CD9-FITC, 5 μl CD81-PE, 5 μl CD45-PerCP, and 3 μl

<table>
<thead>
<tr>
<th>Patient</th>
<th>Diagnosis</th>
<th>Stage</th>
<th>Age (months)</th>
<th>Sex</th>
<th>Sample</th>
<th>Time point</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>NB</td>
<td>4</td>
<td>87</td>
<td>M</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>B</td>
<td>NB</td>
<td>3</td>
<td>13</td>
<td>F</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>C</td>
<td>NB</td>
<td>1</td>
<td>27</td>
<td>M</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>D</td>
<td>NB</td>
<td>4</td>
<td>93</td>
<td>M</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>E</td>
<td>NB</td>
<td>4</td>
<td>22</td>
<td>M</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>F</td>
<td>NB</td>
<td>4S</td>
<td>9</td>
<td>F</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>G</td>
<td>NB</td>
<td>3</td>
<td>160</td>
<td>F</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>H</td>
<td>NB</td>
<td>4</td>
<td>267</td>
<td>M</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>I</td>
<td>NB</td>
<td>2</td>
<td>36</td>
<td>F</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>J</td>
<td>NB</td>
<td>1</td>
<td>29</td>
<td>M</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>K</td>
<td>NB</td>
<td>1</td>
<td>32</td>
<td>F</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>L</td>
<td>NB</td>
<td>4</td>
<td>49</td>
<td>F</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>M</td>
<td>NB</td>
<td>4S</td>
<td>31</td>
<td>F</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>N</td>
<td>NB</td>
<td>4</td>
<td>74</td>
<td>M</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>O</td>
<td>NB</td>
<td>3</td>
<td>34</td>
<td>M</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>P</td>
<td>NB</td>
<td>3</td>
<td>21</td>
<td>F</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>Q</td>
<td>NB</td>
<td>4</td>
<td>53</td>
<td>M</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>R</td>
<td>NB</td>
<td>3</td>
<td>44</td>
<td>F</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>S</td>
<td>NB</td>
<td>4S</td>
<td>43</td>
<td>F</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>T</td>
<td>NB</td>
<td>4</td>
<td>62</td>
<td>M</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>U</td>
<td>NB</td>
<td>4</td>
<td>71</td>
<td>F</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>V</td>
<td>NB</td>
<td>4</td>
<td>106</td>
<td>F</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>W</td>
<td>DIG</td>
<td>NA</td>
<td>25</td>
<td>F</td>
<td>Biopsy</td>
<td>11</td>
</tr>
<tr>
<td>X</td>
<td>GN</td>
<td>NA</td>
<td>179</td>
<td>F</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>Y</td>
<td>Wilms</td>
<td>NA</td>
<td>55</td>
<td>F</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>Z</td>
<td>LS</td>
<td>NA</td>
<td>17</td>
<td>M</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>AA</td>
<td>RM</td>
<td>NA</td>
<td>103</td>
<td>M</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
</tbody>
</table>

**Table 1**

Patient Characteristics

<table>
<thead>
<tr>
<th>Patient</th>
<th>Diagnosis</th>
<th>Age (months)</th>
<th>Sex</th>
<th>Sample</th>
<th>Time point</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>NB</td>
<td>87</td>
<td>M</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>B</td>
<td>NB</td>
<td>13</td>
<td>F</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>C</td>
<td>NB</td>
<td>27</td>
<td>M</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>D</td>
<td>NB</td>
<td>93</td>
<td>M</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>E</td>
<td>NB</td>
<td>22</td>
<td>M</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>F</td>
<td>NB</td>
<td>9</td>
<td>F</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>G</td>
<td>NB</td>
<td>160</td>
<td>F</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>H</td>
<td>NB</td>
<td>267</td>
<td>M</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>I</td>
<td>NB</td>
<td>36</td>
<td>F</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>J</td>
<td>NB</td>
<td>29</td>
<td>M</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>K</td>
<td>NB</td>
<td>32</td>
<td>F</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>L</td>
<td>NB</td>
<td>49</td>
<td>F</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>M</td>
<td>NB</td>
<td>31</td>
<td>F</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>N</td>
<td>NB</td>
<td>74</td>
<td>M</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>O</td>
<td>NB</td>
<td>34</td>
<td>M</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>P</td>
<td>NB</td>
<td>21</td>
<td>F</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>Q</td>
<td>NB</td>
<td>53</td>
<td>M</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>R</td>
<td>NB</td>
<td>44</td>
<td>F</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>S</td>
<td>NB</td>
<td>43</td>
<td>F</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>T</td>
<td>NB</td>
<td>62</td>
<td>M</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>U</td>
<td>NB</td>
<td>71</td>
<td>F</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>V</td>
<td>NB</td>
<td>106</td>
<td>F</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>W</td>
<td>DIG</td>
<td>25</td>
<td>F</td>
<td>Biopsy</td>
<td>11</td>
</tr>
<tr>
<td>X</td>
<td>GN</td>
<td>179</td>
<td>F</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>Y</td>
<td>Wilms</td>
<td>55</td>
<td>F</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>Z</td>
<td>LS</td>
<td>17</td>
<td>M</td>
<td>Biopsy</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>AA</td>
<td>RM</td>
<td>103</td>
<td>M</td>
<td>BM</td>
<td>Diagnosis</td>
</tr>
</tbody>
</table>

**NB**, neuroblastoma; DIG, desmoplastic infantile ganglioglioma; GN, ganglioneuroma; Wilms, Wilms tumor; LS, lung sequester; RM, rhabdomyosarcoma; M, male; F, female; BM, bone marrow; NA, not applicable. PBSC, peripheral blood stem cell.

**b** Weeks after diagnosis.
CD56-APC. In addition, 5 × 10^5 cells were incubated with 2 μl of the unlabeled anti-GD2 antibody, which was diluted 1:50 in phosphate buffered saline (PBS). The volumes of the different antibodies were titrated against the CLB-GA cell line.

The cells were washed twice with PBS. The samples incubated with the isotypic control antibody or the anti-GD2 antibody were incubated for another 30 min with a goat F(ab')2 anti-mouse immunoglobulin G antibody labeled with FITC (1 μl; Caltag Laboratories, Burlingame, UK). After another wash with PBS, we incubated the cells, which had been incubated with the anti-GD2 antibody, with 5 μl CD81-PE, 5 μl CD45-PerCP, and 5 μl CD56-APC. The cells were washed twice and resuspended in PBS. The fluorescence was analyzed on a FACSort flow cytometer (BD Biosystems). At least 10,000 events were counted and analyzed with Cell Quest software (BD Biosystems). NB cells were positive for CD2, CD9, CD81, and CD56 but negative for CD45. Therefore, percentages of CD9^+/CD81^-/CD45^+/CD56^+ and CD2^-/CD81^-/CD45^+/CD56^+ cells were calculated by using the following gating strategy. On a forward scatter/side scatter dot plot, the mononuclear cells were identified by region R1. In all other dot plots, only the cells lying in R1 were represented. Region R2 defined the CD56^+/CD45^- cells. The anti-GD2^+/CD81^- or CD9^+/CD81^- cells were identified by region R3. Samples were considered positive when at least 10 to 20 cells clustering together in R1, R2, and R3 were detected (30).

**Spiking Experiments**

The sensitivity of the four-color flow cytometric and the immunocytochemical assays was evaluated by using spiking experiments. Cells from NB cell line CLB-GA were added to peripheral blood mononuclear cells that had been isolated from healthy volunteers. The concentrations of the NB cells in this dilution experiment were 1%, 0.1%, 0.01%, 0.001%, and 0.0001%. The cells were suspended in RPMI medium.

The dilutions from 1% to 0.001% were evaluated flow cytometrically. With the different dilutions, 3 × 10^5, 5 × 10^5, 3 × 10^4, and 2 × 10^4 events were acquired. The three highest dilutions (0.01% to 0.0001%) were immunocytochemically analyzed, and 5 × 10^5, 1 × 10^5, and 3 × 10^4 cells were evaluated. Because convincing results were obtained, the experiments were performed only once.

**Immunocytochemistry**

After air drying and fixation with 4% paraformaldehyde, large-diameter (17 mm) cytospins (Hettich Zentrifugen, Tuttlingen, Germany) containing 5 × 10^3 mononuclear cells were incubated with 30 μl of an unlabeled monoclonal antibody against GD2 disialoganglioside (clone 14G2a, diluted 1:100 in PBS) in a humidifier at room temperature for 30 min. Controls, in which the primary antibody was replaced by PBS buffer, were included to evaluate the background staining. During a second and a third incubation step, cells were incubated for 30 min with 30 μl of a polyclonal goat anti-mouse antibody (diluted 1:20 in PBS; Dako Corporation) and 30 μl of a polyclonal rabbit anti-GD2 antibody (diluted 1:20 in PBS; Dako Corporation), respectively. Both antibodies were labeled with alkaline phosphatase (31). The bound alkaline phosphatase complexes were stained with the new fuchsin chromogen (Sigma, St. Louis, MO, USA) (32). The slides were counterstained with hematoxylin (Sigma) and examined under a light microscope by two independent observers. Preferably six slides, each containing 5 × 10^5 cells, were analyzed. Cells with a round nucleus, a granular chromatin structure, a sparse amount of cytoplasm, and a strong staining around the entire cell were considered positive. Cells with an aberrant morphology or a pink staining restricted to a subcellular compartment were considered negative. The number of positive single cells and positive cells being part of a Homer-Wright rosette or cell clump were counted. The number of evaluated mononuclear cells was also reported. When the two local observers disagreed or fewer than 10 positive cells were found, pictures of the doubtful cells were sent to an external expert. The results were discussed until agreement was reached.

**Statistics**

After classifying the results in a frequency table, a χ² test was performed to study the relation between the different categories. In addition, a κ test was used to measure the agreement between the different assays.

**RESULTS**

**Phenotype of NB Cells**

Cells from the NB cell line CLB-GA were phenotyped with the four-color flow cytometric assay. The CLB-GA cells showed a strong expression for GD2, CD9, CD81, and CD56 and were negative for CD45 (Fig. 1).

The same markers were analyzed in NB cells from 12 tumor samples. All samples were positive for CD9, CD81, and CD56 and negative for CD45. In one sample, the NB cells lacked GD2 expression. This result was confirmed by immunocytochemistry.

Ten BM samples from healthy donors were included as negative controls. None of these samples scored positive on the immunocytochemical or flow cytometric tests.

**Specificity of the Four-Color Flow Cytometric Assay**

We assessed the specificity of the four-color flow cytometric tests by analyzing tumor samples from a DIG, a ganglioneuroma, a Wilms tumor, a lung sequester, and a rhabdomyosarcoma.

The cells from the DIG were CD9^+/CD81^+/CD56^+/CD45^- and GD2^+/CD81^+/CD56^+/CD45^- and CD2^+/CD81^+/CD56^+/CD45^- Consequently, it was impossible to distinguish between a NB and a DIG by using our four-color flow cytometry assay. The immunocytochemical test was negative because the DIG cells did not meet the morphologic criteria.

Based on our limited results, it was possible to differentiate between NB and the other tumors by using the four-color flow cytometric assay. Tumor cells from the
FIG. 1. Phenotype of the neuroblastoma cell line CLB-GA. The CLB-GA cells show strong GD2, CD9, CD81, and CD56 expressions and are negative for CD45 (immunoglobulin G2a [IGG2a], isotypic control; 14G2a, anti-GD2 monoclonal antibody).
ganglioneuroma and the lung sequester showed weak CD56 and GD2 expressions. The Wilms tumor was negative for CD56 and showed only weak CD9 and CD81 expressions. Rhabdomyosarcoma cells were negative for GD2. None of these tumors scored positive for the immunocytochemical assay.

Sensitivity Experiments

The sensitivity of the flow cytometric assays was evaluated with spiking experiments. CLB-4GA cells were diluted into normal mononuclear cells that had been isolated from peripheral blood taken from healthy volunteers. Four different dilutions (1%, 0.1%, 0.01%, and 0.001%) were analyzed. As few as one NB cell in $10^4$ to $10^5$ normal mononuclear cells could be detected with the four-color flow cytometric assays. Regression analysis was performed:

$$\text{CD9/CD81/CD45/CD56: log (Y) = 0.41 + 1.09 \log (X)}$$

standard error of the slope = 0.10, corresponding

$$P = 0.0082, \text{residual standard deviation} = 0.22$$

$$\text{GD2/CD81/CD45/CD56: log (Y) = 0.44 + 1.23 \log (X)}$$

Both tests performed equally well during the spiking experiment.

It is important to analyze a sufficient number of cells when the tumor load in a sample is low (<0.01%) because samples can be considered positive only when at least 10 to 20 cells with the correct immunophenotype, clustering together in the different dot plots, are detected. In our study, at least $2 \times 10^4$ cells were flow cytometrically analyzed. This means that a sensitivity of 0.1% to 0.2% was reached.

The sensitivity of the immunocytochemical test was also evaluated with spiking experiments. Three different dilutions (0.01%, 0.001%, and 0.0001%) were analyzed. When $3 \times 10^3$ cells were evaluated, one NB cell in $10^5$ normal mononuclear cells could be detected.

Detection of NB Cells in BM and PBSC Preparations

Twenty-eight BM samples and three PBSC preparations were analyzed with the flow cytometric assay. The results are summarized in Table 2. Flow cytometry scored positive in four BM samples, and 24 samples were negative.

### Table 2

<table>
<thead>
<tr>
<th>Patients</th>
<th>Samples</th>
<th>Stage</th>
<th>Time point*</th>
<th>Flow Cytometry</th>
<th>Cells immunocytochemistry</th>
<th>Cells evaluated</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>BM 1</td>
<td>4</td>
<td>Diagnosis</td>
<td></td>
<td>2.00E + 04</td>
<td>1.00E + 06</td>
</tr>
<tr>
<td>B</td>
<td>BM 2</td>
<td>3</td>
<td>Diagnosis</td>
<td></td>
<td>2.00E + 04</td>
<td>5.00E + 05</td>
</tr>
<tr>
<td></td>
<td>BM 3</td>
<td>3</td>
<td>Diagnosis</td>
<td></td>
<td>2.00E + 04</td>
<td>7.00E + 05</td>
</tr>
<tr>
<td>E</td>
<td>BM 4</td>
<td>2</td>
<td>Diagnosis</td>
<td></td>
<td>2.00E + 04</td>
<td>1.70E + 06</td>
</tr>
<tr>
<td>M</td>
<td>BM 5 2</td>
<td>4S</td>
<td>Diagnosis</td>
<td></td>
<td>2.00E + 04</td>
<td>1.00E + 06</td>
</tr>
<tr>
<td>N</td>
<td>BM 7</td>
<td>4</td>
<td>Diagnosis</td>
<td></td>
<td>2.00E + 04</td>
<td>3.00E + 06</td>
</tr>
<tr>
<td></td>
<td>BM 8</td>
<td>4</td>
<td>+</td>
<td></td>
<td>2.00E + 04</td>
<td>2.50E + 06</td>
</tr>
<tr>
<td></td>
<td>BM 9</td>
<td>4</td>
<td>+</td>
<td>3.38E + 05</td>
<td>+</td>
<td>1.50E + 06</td>
</tr>
<tr>
<td>O</td>
<td>BM 10</td>
<td>3</td>
<td>Diagnosis</td>
<td></td>
<td>2.00E + 04</td>
<td>3.00E + 06</td>
</tr>
<tr>
<td>P</td>
<td>BM 11</td>
<td>3</td>
<td>Diagnosis</td>
<td></td>
<td>2.00E + 04</td>
<td>3.00E + 05</td>
</tr>
<tr>
<td>Q</td>
<td>BM 12</td>
<td>4</td>
<td>Diagnosis</td>
<td></td>
<td>2.00E + 04</td>
<td>1.50E + 06</td>
</tr>
<tr>
<td>R</td>
<td>BM 13</td>
<td>3</td>
<td>Diagnosis</td>
<td></td>
<td>2.00E + 04</td>
<td>5.00E + 05</td>
</tr>
<tr>
<td>S</td>
<td>BM 14</td>
<td>3</td>
<td>Relapse</td>
<td></td>
<td>2.00E + 04</td>
<td>1.50E + 06</td>
</tr>
<tr>
<td>T</td>
<td>BM 15</td>
<td>3</td>
<td>Relapse</td>
<td></td>
<td>2.00E + 04</td>
<td>1.50E + 06</td>
</tr>
<tr>
<td>U</td>
<td>BM 16</td>
<td>4S</td>
<td>Diagnosis</td>
<td></td>
<td>2.00E + 04</td>
<td>4.00E + 05</td>
</tr>
<tr>
<td>V</td>
<td>BM 17</td>
<td>4</td>
<td>13</td>
<td></td>
<td>2.00E + 04</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>BM 18</td>
<td>4</td>
<td>23</td>
<td></td>
<td>2.00E + 04</td>
<td>7.00E + 05</td>
</tr>
<tr>
<td></td>
<td>BM 19</td>
<td>4</td>
<td>23</td>
<td></td>
<td>2.00E + 04</td>
<td>1.20E + 06</td>
</tr>
<tr>
<td></td>
<td>BM 20</td>
<td>4</td>
<td>40</td>
<td></td>
<td>2.00E + 04</td>
<td>9.00E + 05</td>
</tr>
<tr>
<td></td>
<td>BM 21</td>
<td>4</td>
<td>Diagnosis</td>
<td></td>
<td>2.00E + 04</td>
<td>7.00E + 05</td>
</tr>
<tr>
<td></td>
<td>BM 22</td>
<td>4</td>
<td>11</td>
<td></td>
<td>2.00E + 04</td>
<td>4.00E + 05</td>
</tr>
<tr>
<td></td>
<td>BM 23</td>
<td>4</td>
<td>Diagnosis</td>
<td></td>
<td>2.00E + 04</td>
<td>1.40E + 05</td>
</tr>
<tr>
<td></td>
<td>BM 24</td>
<td>4</td>
<td>Diagnosis</td>
<td></td>
<td>2.00E + 04</td>
<td>1.00E + 05</td>
</tr>
<tr>
<td></td>
<td>BM 25</td>
<td>4</td>
<td>7</td>
<td></td>
<td>2.00E + 04</td>
<td>3.00E + 06</td>
</tr>
<tr>
<td></td>
<td>BM 26</td>
<td>4</td>
<td>7</td>
<td></td>
<td>2.00E + 04</td>
<td>3.00E + 06</td>
</tr>
<tr>
<td></td>
<td>BM 27</td>
<td>4</td>
<td>13</td>
<td></td>
<td>2.00E + 04</td>
<td>2.00E + 05</td>
</tr>
<tr>
<td></td>
<td>BM 28</td>
<td>4</td>
<td>13</td>
<td></td>
<td>2.00E + 04</td>
<td>9.50E + 05</td>
</tr>
<tr>
<td></td>
<td>PBSC 1</td>
<td>4</td>
<td>13</td>
<td></td>
<td>1.00E + 05</td>
<td>3.00E + 06</td>
</tr>
<tr>
<td></td>
<td>PBSC 2</td>
<td>4</td>
<td>13</td>
<td></td>
<td>4.00E + 02</td>
<td>2.50E + 06</td>
</tr>
<tr>
<td></td>
<td>PBSC 3</td>
<td>4</td>
<td>13</td>
<td></td>
<td>2.00E + 04</td>
<td>1.00E + 06</td>
</tr>
</tbody>
</table>

*Weeks after diagnosis, BM, bone marrow; PBSC, peripheral blood stem cell.
These results were compared with the results of an immunocytochemical assay. We found concordant results in 24 of 28 BM samples (86%). The four samples were positive for the flow cytometric assay and for the immunocytochemical test. Twenty BM samples were negative for both assays. During therapy, discordant results were found in four BM samples and two PBSC preparations. Although these samples scored positive for the immunocytochemical test, fewer than 10 cells clustering together in the different dot plots were found after four-color flow cytometric analysis. In these samples, the number of residual NB cells was very small (Table 3). In addition, there was a significant difference in the number of mononuclear cells analyzed by both techniques. Using flow cytometry, an average of 1.2 × 10^6 cells was analyzed. In contrast, an average of 1.2 × 10^4 cells was evaluated using immunocytochemistry.

Figure 2 shows an example of a BM sample that scored positive for the flow cytometric and immunocytochemical assays. Figure 3 shows an example of discordant results. Three positive cells were detected after immunocytochemical analysis of 3 × 10^6 mononuclear cells. Because only 5 × 10^5 cells were flow cytometrically analyzed, theoretically only 0.5 tumor cells could have been detected by the flow cytometric assays. Both samples were only recently analyzed and are not included in the statistically processed data.

There was no change in the intensity of the antigen expression or the immunophenotype of the tumor cells when, on more than one occasion, NB cells were detected.

Based on the results of the BM and PBSC samples, an association between the flow cytometric and immunocytochemical test was found ($\chi^2 = 6.4, P = 0.011, df = 1$). In addition, the $\kappa$ statistic showed a moderate level of agreement between tests ($\kappa = 0.48$).

**DISCUSSION**

Many children with NB present with disseminated disease, including BM involvement. The detection of contaminating NB cells in the BM plays an important role in clinical staging and risk assessment. Moreover, the detection of disseminating tumor cells in the autologous stem cell preparations may signify the need for ex vivo manipulations such as tumor cell purging and/or CD34^+ cell selection to avoid the reinfusion of malignant cells (33).

Different techniques for the detection of residual NB cells such as immunocytochemistry, RT-PCR, and flow cytometry are described in the literature (12–19).

Flow cytometry has been used in the clinical practice of hematology and oncology for many years and has many advantages compared with immunocytochemistry or RT-PCR. The flow cytometric assay provides information on the antigen expression from a single tumor cell in a multicolor assay and allows the analysis of large cell numbers. In addition, flow cytometry, like automated fluorescence microscopy, allows a more objective quantification of the number of NB cells compared with conventional immunocytochemistry or quantitative RT-PCR (17,34).

In this study, we developed a new four-color flow cytometric assay using CD9, CD81, CD45, and CD56 to detect NB cells in BM. A second four-color assay based on the use of anti-GD2, CD81, CD45, and CD56 was evaluated (18). Twelve NB tumor samples were analyzed. One tumor lacked GD2 expression. This result was confirmed by the immunocytochemical assay. The sample was taken from a 13-year-old girl with a palpable mass in the abdomen. TheNSE serum concentration and the urinary dopamine concentration were elevated. Magnetic resonance imaging showed a large tumor extending from the right kidney to the pelvis and into the lumbar foramina. No metastases were found. Immunohistologically, the tumor cells showed strong CD56 and NSE expressions. They were moderately positive for chromogranin and synaptophysin. There were no obvious signs of cytodifferentiation. After genetic analysis, no MYCN amplification or 1p deletion was found, but 70% of the tumor cells had an extra copy of chromosome 17 and 18% showed 17q gain. The diagnosis of a stroma-poor, poorly differentiated, stage 3 NB was made for this very unusual GD2-negative case. Because nearly all NBs express high levels of GD2 (21–23), immunotherapy using anti-GD2 antibodies is considered a promising treatment strategy (35). However, when the primary tumor lacks GD2 expression, treatment with anti-GD2 immunotherapy is useless. Therefore, we recommend the evaluation of GD2 expression in tumor samples from all NB patients at diagnosis.

It can be difficult to differentiate between NB and other members of the small blue round cell tumor family. Our results suggest that the four-color flow cytometric assays are helpful in making a differential diagnosis. Only cells

<table>
<thead>
<tr>
<th>Patients</th>
<th>Samples</th>
<th>Stage</th>
<th>Time point</th>
<th>Flow cytometry</th>
<th>Cells evaluated</th>
<th>Immunochemistry</th>
<th>Cells evaluated</th>
<th>No. of positive cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>BM 17</td>
<td>4</td>
<td>13</td>
<td>−</td>
<td>2.00E + 04</td>
<td>+</td>
<td>1.00E + 06</td>
<td>5</td>
</tr>
<tr>
<td>V</td>
<td>BM 20</td>
<td>4</td>
<td>40</td>
<td>−</td>
<td>2.00E + 04</td>
<td>+</td>
<td>9.00E + 05</td>
<td>4</td>
</tr>
<tr>
<td>N</td>
<td>BM 27</td>
<td>4</td>
<td>13</td>
<td>−</td>
<td>2.00E + 04</td>
<td>+</td>
<td>2.20E + 05</td>
<td>2</td>
</tr>
<tr>
<td>N</td>
<td>BM 28</td>
<td>4</td>
<td>13</td>
<td>−</td>
<td>2.00E + 04</td>
<td>+</td>
<td>9.50E + 05</td>
<td>3</td>
</tr>
<tr>
<td>N</td>
<td>PBSC 1</td>
<td>4</td>
<td>13</td>
<td>−</td>
<td>1.00E + 05</td>
<td>+</td>
<td>3.00E + 06</td>
<td>2</td>
</tr>
<tr>
<td>N</td>
<td>PBSC 3</td>
<td>4</td>
<td>13</td>
<td>−</td>
<td>2.00E + 04</td>
<td>+</td>
<td>1.00E + 06</td>
<td>5</td>
</tr>
</tbody>
</table>

*Weeks after diagnosis, BM, bone marrow; PBSC, peripheral blood stem cell.
FIG. 2. Example of a bone marrow positive for the flow cytometric and immunocytochemical assays. A: Flow cytometric results. On a forward scatter (FSC)/side scatter (SSC) dot plot, the mononuclear cells were identified by region R1. In all other dot plots, only the cells lying in R1 were represented. R2 defined the CD56+/H11001/CD45+/H11002 cells. The CD9+/H11001/CD81+/H11001 cells were identified by R3. Because more than 10 to 20 cells clustering together in R1, R2, and R3 were detected, the sample was considered positive. A total of $5 \times 10^5$ cells/tube was analyzed (immunoglobulin G2a, isotypic control; 14G2a, anti-GD2 monoclonal antibody). B: Immunocytochemical results. The sample was considered positive because clusters of positive cells fulfilling the criteria were detected. A total of $3 \times 10^6$ was immunocytochemically analyzed.
FIG. 3. Example of a bone marrow in which the residual neuroblastoma cells were found only after immunocytochemical analysis. A: Flow cytometric results. No tumor cells were detected. A total of $5 \times 10^5$ cells/tube were analyzed. B: Immunocytochemical results. One single cell and a cluster of three tumor cells fulfilling the criteria were found.
from the DIG tumor, containing mature gangliocytes, were CD9/CD81/CD45/GD2/CD56 and GD2/CD81/CD45/CD56. Because the CD56 expression on the DIG cells was as high as the CD56 expression on NB cells, it was impossible to differentiate between them. The GD2 expression on the ganglioneuroma cells was very weak. Others have reported the absence of GD2 in more differentiated ganglioneuroblastomas and ganglioneuromas (18,20).

In six samples, discordant flow cytometric and immunocytochemical results were obtained. In these samples, taken during therapy, only a few residual NB cells were found after immunocytochemical analysis. The discrepancy was probably due to sample variability and the difference in the number of analyzed cells. Because the conventional immunocytochemical assay was considered the gold standard for minimal residual disease detection in patients with an undifferentiated NB, priority was given to this test and $3 \times 10^6$ cells were immunocytochemically evaluated. The remaining cells were analyzed with the two four-color flow cytometric assays. Because BM samples taken during therapy are often aplastic, in some cases only $2 \times 10^4$ cells could be flow cytometrically analyzed. This means a sensitivity of 0.1% was reached. Seeger et al. reported that a sensitivity of 0.1% (100 NB cells/$10^5$ BM cells) after three to four cycles of therapy most likely represents the clinically useful threshold for patients (4). Children who demonstrated less sensitivity survived, whereas those with greater sensitivity died. Therefore, the collection of $1 \times 10^5$ cells/tube was considered adequate.

Two combinations of different monoclonal antibodies are currently used to screen for residual NB cells. This means that at least $2 \times 10^6$ cells must be analyzed to reach a sensitivity of 0.001%. This amount of cells is sometimes not available. In the near future, we plan to develop a five-color flow cytometric assay so that we can decrease the amount of required cells by one-third.

The discrepancy between the immunocytochemical and flow cytometric results can also be explained by a difference in sensitivity. Our immunocytochemical test has a targeted sensitivity of 0.0001%. This is approximately 10 to 100 times more sensitive than that reported in previous publications (4,28,36,37). The sensitivity of the assay is not limited by the technique but by the number of analyzed cells. When enough cells are analyzed, a high sensitivity can be reached. The Poisson distribution, $f(X) = \mu^X e^{-\mu}/X!$, can be used to calculate the statistics of tumor detection (28). At least $3 \times 10^6$ cells must be analyzed to have a 95% chance to detect one tumor cell in a million normal mononuclear cells. In previous reports, fewer than $3 \times 10^5$ cells were analyzed, which decreased the sensitivity of the assays. In contrast, the flow cytometric assay has a sensitivity of 0.01% to 0.001% based on spiking experiments. This corresponds with previously published reports (12,14,18). In our study, a sensitivity of 0.1% to 0.2% was reached because at least $2 \times 10^4$ cells were flow cytometrically analyzed.

In conclusion, the sensitivity of the flow cytometric assays is lower than that of the immunocytochemical assay in spiking experiments. However, the flow cytometric assays can be used to screen for residual cells in clinical samples with a sensitivity of one NB cell in $10^4$ to $10^5$ normal mononuclear cells. They are less labor intensive and time consuming than the immunocytochemical test. The results are available only a few hours after collection, and the technique is simple and cost effective. Further, in contrast to our immunocytochemical assay, the four-color flow cytometric tests can be used to screen for residual NB cells in case of a GD2-negative primary tumor. Therefore, we recommend flow cytometry for the detection of residual NB cells in BM.

ACKNOWLEDGMENTS

The authors thank Prof. Dr. F. Speleman for technical assistance and the provision of the NB cell line CLB-GA. They also thank Dr. K. Beiske who reviewed the doubtful cases and helped taking pictures of immunocytochemically stained cells. They are also grateful to D. Claey for skillful assistance in the immunocytochemical and flow
cytometric assays. They also thank Dr. C. De Valck, Dr. N. Francotte, Dr. C. Hoyoux, Dr. E. Michiels, Dr. A. Uyttebroeck and Dr. J. van der Werf for sending us samples of NB patients.

LITERATURE CITED


2.3. Paper 6

ELAVL4: A NEW MOLECULAR MARKER FOR DETECTION OF MINIMAL RESIDUAL DISEASE IN NEUROBLASTOMA USING REAL-TIME QUANTITATIVE RT-PCR

Swerts K., De Moerloose B., Dhooge C., Vandesompele J., Hoyoux C., Benoit Y., Laureys G. and Philippé J.

Clinical Cancer Research (Submitted)
ELAVL4: a new molecular marker for detection of minimal residual disease in neuroblastoma using real-time quantitative RT-PCR

Swerts K., De Moerloose B., Dhooge C., Vandesompele J., Hoyoux C., Benoit Y., Laureys G. and Philippé J.

Clinical Cancer Research (Submitted)

ABSTRACT

Purpose:
Reliable detection of neuroblastoma (NB) cells in bone marrow (BM) is critical since BM involvement influences staging, risk assessment and evaluation of therapeutic response in NB patients. Standard cytomorphological examination of BM aspirates appears not sensitive enough to detect single tumor cells. Therefore, more sensitive and specific detection methods are needed.

Experimental design:
Disseminated NB cells were detected by means of tyrosine hydroxylase (TH), GD2 synthetase (GALGT) and 'Embryonic Lethal, Abnormal Vision, Drosophila-like 4' (ELAVL4) real-time quantitative RT-PCR (QPCR). The sensitivity of the molecular assays was assessed using spiking experiments. Ninety-seven tumor, BM, peripheral blood (PB) or peripheral blood stem cell (PBSC) samples from 30 patients were analyzed. The QPCR results were compared to those of a standardized immunocytochemical (IC) assay.

Results:
The molecular markers were highly expressed in all evaluated tumor samples. In addition, 32%, 11% and 38% of all BM, PB and PBSC samples scored positive for TH, GALGT or ELAVL4, respectively. The sensitivity of the TH and ELAVL4 assay was $1:10^6$. By contrast, one NB cell in $10^4$ mononuclear cells could be detected using GALGT QPCR.
The potential prognostic value of TH, GALGT and ELAVL4 QPCR was assessed by analyzing subsequent samples from 3 stage 4 patients. Preliminary results indicate that persistence of high ELAVL4 expression has prognostic value.

Conclusions:
ELAVL4 QPCR can be used to detect residual NB cells in clinical samples. However, it is advisable to combine several molecular markers (ELAVL4-TH) and screening techniques (IC) to assure the reliable assessment of MRD in NB.
INTRODUCTION

Neuroblastoma (NB), the most common extracranial malignant solid tumor in children, originates from the sympathetic nervous tissue in the neural crest. Its clinical behavior ranges from a localized tumor with a good prognosis to disseminated disease with unfavorable outcome. Many clinical and biological factors have been shown to affect the prognosis of NB patients (1).

The evaluation of BM metastasis is crucial for correct clinical staging and risk assessment at diagnosis. The detection of residual NB cells in BM is also important for monitoring therapeutic response during treatment. In addition, screening of autologous stem cell preparations is crucial since the re-infusion of contaminated stem cell products could lead to systemic recurrence (2-4).

According to the International Neuroblastoma Staging System, conventional cytology of BM smears is still the only accepted technique for the detection of residual NB cells (5). However, the sensitivity of this method is limited since a tumor cell number below 0.1 % cannot be detected by conventional cytomorphology (6, 7). Therefore, the development of more sensitive and specific detection methods is indispensable.

During the last decades, sensitive assays based on immunocytology (8-10), automatic immunofluorescence plus FISH (11) and flow cytometry (12-16) were evaluated. In addition, several molecular tests based on reverse transcriptase-polymerase chain reaction (RT-PCR) have been developed. The usefulness of various tumor specific gene-transcripts such as tyrosine hydroxylase (17-19), GAGE (20, 21), MAGE 1 to 4 (22), neuroendocrine protein PGP 9.5 (23, 24), GD2 synthetase (25, 26), dopamine decarboxylase (27) and chromogranin A (28) was assessed.

Most neuroblastomas (90-95%) are characterized by a high catecholamine production. The first and rate-limiting step in the biosynthesis of catecholamines is catalyzed by tyrosine hydroxylase (TH). Therefore, this enzyme is often used as a molecular marker for the detection of disseminated NB cells in peripheral blood (PB), BM and peripheral blood stem cell (PBSC) preparations.

GD2 synthetase (GALGT) (β1,4-N-acetylgalactosaminyltransferase) catalyzes the transfer of 1,4-N-acetylglactosamine to GD3 ganglioside and plays a key role in GD2 disialoganglioside biosynthesis (29). Since GALGT is strongly expressed in neural crest-derived tumors such as melanoma and neuroblastoma (30), the enzyme is frequently used as a molecular marker for minimal residual disease (MRD) detection in NB.

ELAVL4 (Embryonic Lethal, Abnormal Vision, Drosophila-like 4) or HuD (Hu antigen D) is also a potential molecular marker which can be used to detect rare NB cells. The gene belongs to the elav gene family of Drosophila melanogaster and has been reported to be
highly specific for neuroectodermally derived tumors (31). So far, the use of ELAVL4 mRNA for MRD detection in NB has not been reported.

In this study, real-time quantitative RT-PCR (QPCR) tests for TH, GALGT and ELAVL4 mRNA were developed and evaluated. The sensitivity of the assays was assessed using both spiking experiments and clinical samples. Thirty PB or BM samples from patients without malignant disease were included as negative controls. The TH, GALGT and ELAVL4 gene expression was analyzed in 18 tumor, 70 BM, 5 PB and 4 PBSC samples, collected at diagnosis or during treatment. Finally, the molecular results were compared to those of a standardized anti-GD2 immunocytochemical reference assay.
MATERIALS AND METHODS

Patients and sample preparation
Between April 2001 and October 2004, 18 tumor, 70 BM, 5 PB and 4 PBSC samples from one ganglioneuroblastoma (GNB), one ganglioneuroma (GN) and 28 neuroblastoma patients, taken at diagnosis or during treatment, were examined. Bilateral BM samples were considered as two different samples because they were analyzed separately. Patient characteristics are summarized in Table 1. The MYCN and 1p status was evaluated using FISH following the guidelines of the European Neuroblastoma Quality Assessment (ENQUA) group (32).

The NB patients were diagnosed and staged according to the International Neuroblastoma Staging System (1). Treatment depended on the patient’s age, tumor stage and biological risk factors. The Ethical Committee approved the study and informed consent was obtained from the patients and/or parents.

From fresh tumor samples, single cell suspensions were prepared using collagenase (1000 U/ml in RPMI for 1h at 37 °C). Subsequently, cells were washed and cell clumps were removed using a 70 µm nylon membrane filter.

BM, PB and PBSC samples were collected into EDTA containing tubes. Immediately after collection, mononuclear cells were isolated using density gradient centrifugation on Fycoll-Hypaque.

All cells were resuspended in 75% ethanol and stored at -80 °C until analysis.

Control Samples
The NB cell line CLB-GA was included in the experiments as positive control. BM and PB samples from 30 adults without malignant disease were used as negative controls. BM samples from healthy pediatric patients were not included in the study because we expected it to be too difficult to obtain informed consent from their parents.

Sensitivity
Spiking experiments were performed to evaluate the sensitivity of the TH, GALGT and ELAVL4 QPCR assays. RNA from the NB cell line IMR32 was added to 1 µg of RNA isolated from normal PB mononuclear cells. Five 10-fold dilutions containing 10,000 to 10 pg of IMR32 RNA were analyzed.

The sensitivity of the molecular assays was also assessed using patient samples. QPCR results were compared to those of an immunocytochemical reference assay (sensitivity 1:10^6) which was considered to be the gold standard. All samples were handled as described below.
Table 1: Patient characteristics (Dx: Diagnosis; M: Male; F: Female; MYCN +: MYCN amplification; 1p del: Deletion of 1p; NE: not evaluated; NB: neuroblastoma; GNB: ganglioneuroblastoma; GN: ganglioneuroma)

<table>
<thead>
<tr>
<th>Nr</th>
<th>Diagnosis</th>
<th>Stage</th>
<th>Age at Dx (months)</th>
<th>Sex</th>
<th>MYCN +</th>
<th>1p del</th>
<th>Catecholamine production at Dx</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NB</td>
<td>1</td>
<td>2</td>
<td>M</td>
<td>NO</td>
<td>NO</td>
<td>Elevated</td>
</tr>
<tr>
<td>2</td>
<td>NB</td>
<td>2a</td>
<td>0</td>
<td>F</td>
<td>NO</td>
<td>IMBALANCE</td>
<td>Elevated</td>
</tr>
<tr>
<td>4</td>
<td>NB</td>
<td>3</td>
<td>27</td>
<td>M</td>
<td>NO</td>
<td>NO</td>
<td>Elevated</td>
</tr>
<tr>
<td>6</td>
<td>NB</td>
<td>3</td>
<td>14</td>
<td>M</td>
<td>NO</td>
<td>NO</td>
<td>Elevated</td>
</tr>
<tr>
<td>7</td>
<td>NB</td>
<td>3</td>
<td>9</td>
<td>F</td>
<td>NO</td>
<td>NO</td>
<td>NE</td>
</tr>
<tr>
<td>8</td>
<td>NB</td>
<td>3</td>
<td>157</td>
<td>F</td>
<td>NO</td>
<td>IMBALANCE</td>
<td>Elevated</td>
</tr>
<tr>
<td>9</td>
<td>NB</td>
<td>3</td>
<td>6</td>
<td>F</td>
<td>NO</td>
<td>NO</td>
<td>Elevated</td>
</tr>
<tr>
<td>10</td>
<td>NB</td>
<td>3</td>
<td>8</td>
<td>F</td>
<td>NO</td>
<td>YES</td>
<td>Elevated</td>
</tr>
<tr>
<td>11</td>
<td>NB</td>
<td>3</td>
<td>23</td>
<td>F</td>
<td>NO</td>
<td>IMBALANCE</td>
<td>Elevated</td>
</tr>
<tr>
<td>12</td>
<td>NB</td>
<td>4</td>
<td>82</td>
<td>M</td>
<td>NO</td>
<td>IMBALANCE</td>
<td>Elevated</td>
</tr>
<tr>
<td>13</td>
<td>NB</td>
<td>4</td>
<td>25</td>
<td>F</td>
<td>YES</td>
<td>NO</td>
<td>Not Elevated</td>
</tr>
<tr>
<td>14</td>
<td>NB</td>
<td>4</td>
<td>10</td>
<td>F</td>
<td>YES</td>
<td>NO</td>
<td>Elevated</td>
</tr>
<tr>
<td>15</td>
<td>NB</td>
<td>4</td>
<td>72</td>
<td>M</td>
<td>NO</td>
<td>NO</td>
<td>Elevated</td>
</tr>
<tr>
<td>16</td>
<td>NB</td>
<td>4</td>
<td>16</td>
<td>F</td>
<td>YES</td>
<td>YES</td>
<td>Elevated</td>
</tr>
<tr>
<td>17</td>
<td>NB</td>
<td>4</td>
<td>53</td>
<td>M</td>
<td>NO</td>
<td>NO</td>
<td>Elevated</td>
</tr>
<tr>
<td>18</td>
<td>NB</td>
<td>4</td>
<td>3</td>
<td>M</td>
<td>NO</td>
<td>NO</td>
<td>Elevated</td>
</tr>
<tr>
<td>19</td>
<td>NB</td>
<td>4</td>
<td>70</td>
<td>M</td>
<td>NO</td>
<td>YES</td>
<td>Elevated</td>
</tr>
<tr>
<td>20</td>
<td>NB</td>
<td>4</td>
<td>26</td>
<td>M</td>
<td>NO</td>
<td>NO</td>
<td>Elevated</td>
</tr>
<tr>
<td>21</td>
<td>NB</td>
<td>4</td>
<td>26</td>
<td>M</td>
<td>YES</td>
<td>NO</td>
<td>Not Elevated</td>
</tr>
<tr>
<td>22</td>
<td>NB</td>
<td>4</td>
<td>21</td>
<td>F</td>
<td>NO</td>
<td>NO</td>
<td>Elevated</td>
</tr>
<tr>
<td>23</td>
<td>NB</td>
<td>4</td>
<td>70</td>
<td>F</td>
<td>NO</td>
<td>NO</td>
<td>Elevated</td>
</tr>
<tr>
<td>24</td>
<td>NB</td>
<td>4</td>
<td>31</td>
<td>M</td>
<td>NO</td>
<td>NO</td>
<td>Elevated</td>
</tr>
<tr>
<td>25</td>
<td>NB</td>
<td>4</td>
<td>35</td>
<td>F</td>
<td>YES</td>
<td>YES</td>
<td>Elevated</td>
</tr>
<tr>
<td>26</td>
<td>NB</td>
<td>4</td>
<td>6</td>
<td>F</td>
<td>NO</td>
<td>NO</td>
<td>Elevated</td>
</tr>
<tr>
<td>27</td>
<td>NB</td>
<td>4</td>
<td>92</td>
<td>F</td>
<td>YES</td>
<td>YES</td>
<td>Elevated</td>
</tr>
<tr>
<td>28</td>
<td>NB</td>
<td>4S</td>
<td>4</td>
<td>F</td>
<td>NO</td>
<td>NO</td>
<td>Not Elevated</td>
</tr>
<tr>
<td>29</td>
<td>GNB</td>
<td>2</td>
<td>33</td>
<td>F</td>
<td>NO</td>
<td>NO</td>
<td>Elevated</td>
</tr>
<tr>
<td>30</td>
<td>GN</td>
<td>3</td>
<td>44</td>
<td>F</td>
<td>NO</td>
<td>NO</td>
<td>Elevated</td>
</tr>
</tbody>
</table>

RNA isolation and Reverse Transcription

Total RNA was extracted from 5x10⁶ to 1x10⁷ cells using TRIzol Reagent (Invitrogen, Carlsbad, CA, USA) following the instructions of the manufacturer. Briefly, TRIzol Reagent maintains the integrity of the RNA while disrupting cells and dissolving cell components. The addition of chloroform followed by centrifugation separates the solution into an aqueous and an organic phase. The RNA remains exclusively in the aqueous phase and is recovered by precipitation with isopropyl alcohol. Finally, the RNA is resuspended in RNase-free water. The concentration and purity of the recovered RNA were measured by OD at 260 and 280 nm. The integrity of the isolated RNA was verified by quantitative real-time RT-PCR for human beta-2-microglobulin and ubiquitin C.
Total RNA (1 µg in 5 µl) was denatured at 95 °C for 5 min (OmniGene thermal cycler, Thermo Electron, Waltham, MA, USA) and subsequently placed on ice. An equal volume of the RT reaction mixture (5 µl) was added. The final concentrations of the reagents in the reaction (10 µl) were 1x TaqMan Buffer A (Applied Biosystems, Foster City, CA, USA), 1 mM of each deoxynucleotide triphosphate (Amersham Biosciences, Piscataway, NJ, USA), 8 mM MgCl₂ (Sigma-Aldrich, St. Louis, MO, USA), 0.3 µg of Random Hexamer Primers (Invitrogen), 8 units of RNA Guard (Amersham Biosciences) and 10 units of Moloney Murine Leukemia Virus reverse transcriptase (Amersham Biosciences). The RNA was reverse transcribed at 37 °C for 1 h followed by the inactivation of the enzyme at 95 °C for 5 min.

**Real-time quantitative RT-PCR**

Detection chemistry:

TH, GALGT and ELAVL4 transcripts were quantified using the ABI PRISM 7700 Sequence Detection System (Applied Biosystems). A fluorogenic probe, labeled with a reporter (FAM) at the 5’ end and a quencher (TAMRA) at the 3’ end, was used. All primers and probes were designed using the Primer Express software (Applied Biosystems) and are listed in Table 2. To avoid the amplification of genomic DNA, primers were located in successive exons. The primers and probe for TH were designed to hybridize to exons not influenced by alternative splicing.

The quencher suppresses the fluorescent emission of the reporter whenever the probe is intact. During the extension phase of the PCR, the annealed probe is cleaved by the 5’ exonuclease activity of the Taq polymerase resulting in the release of the reporter. This is coupled to an increase in the fluorescent emission. Since this event occurs during every PCR cycle, the fluorescence increases proportional to the amount of synthesized PCR products. After a number of PCR cycles, the fluorescence will reach a threshold. The position of the threshold is chosen based on the variability of the baseline data. The point at which the fluorescence crosses the threshold is defined as Ct value.

Quantification:

In order to construct relative standard curves for the quantification of TH, GALGT and ELAVL4 transcripts, serial dilutions of cDNA from the NB cell line CLB-GA were used. For each assay, 10-fold serial dilutions, starting from relative copy number 100,000 (corresponding to 100 ng of CLB-GA RNA) to relative copy number 10 (corresponding to 10 pg of CLB-GA RNA), were analyzed in duplicate. A standard curve was established by plotting the Ct value versus the logarithm of the relative copy number (Figure 1A).
Table 2: Sequences and locations of the primers and probes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward primer</th>
<th>Primer and probe</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>TH</td>
<td>5' ATT GCT GAG ATC GCC TTC CA 3'</td>
<td>5' (FAM) ACA GCC AGC GCC ACC CGA TTC (TAMRA) 3'</td>
<td>Exon 6</td>
</tr>
<tr>
<td>NM_00360</td>
<td>5' AAT CTC GGC GGT GTA CTC 3'</td>
<td>Exon 7</td>
<td></td>
</tr>
<tr>
<td>GALGT</td>
<td>5' AGC CGA AGC TAC CAG ACC AA 3'</td>
<td>Exon 6</td>
<td></td>
</tr>
<tr>
<td>NM_001478</td>
<td>5' GGA TAG TGA AAG CAG CCT GT 3'</td>
<td>Exon 7</td>
<td></td>
</tr>
<tr>
<td>Probe</td>
<td>5' (FAM) ACA GCA GAC ACA GTG ACC CGA TTC ACC (TAMRA) 3'</td>
<td>Exon 6-7</td>
<td></td>
</tr>
<tr>
<td>ELAVL4</td>
<td>5' AAT GGA CTC AGA CTC CAG ACC AA 3'</td>
<td>Exon 6-7</td>
<td></td>
</tr>
<tr>
<td>NM_021952</td>
<td>5' CAT AGA GGT TAG CAT CAC TGA TGG A 3'</td>
<td>Exon 6</td>
<td></td>
</tr>
<tr>
<td>Probe</td>
<td>5' (FAM) ACA GCA GAC ACA GTG ACC CGA TTC ACC (TAMRA) 3'</td>
<td>Exon 5-6</td>
<td></td>
</tr>
<tr>
<td>B2M</td>
<td>5' GAG TAT GCC TGC CGT GTG 3'</td>
<td>Exon 2</td>
<td></td>
</tr>
<tr>
<td>NM_004048</td>
<td>5' AAT CCA AAT GCG GCA GCT 3'</td>
<td>Exon 4</td>
<td></td>
</tr>
<tr>
<td>Probe</td>
<td>5' (FAM) CCT CCA TGA TGC TGC TTA CAT GTC TC (TAMRA) 3'</td>
<td>Exon 3-4</td>
<td></td>
</tr>
<tr>
<td>UBC</td>
<td>5' TGT GAT CGT CAC TAC ATG CGC 3'</td>
<td>Exon 2-1</td>
<td></td>
</tr>
<tr>
<td>NM_021009</td>
<td>5' CGA TGG TGG TAC TGC CTG C 3'</td>
<td>Exon 2</td>
<td></td>
</tr>
<tr>
<td>Probe</td>
<td>5' (FAM) AAG ACT CTG ACT GGT AGG ACC ATC ACC CTC GA (TAMRA) 3'</td>
<td>Exon 2</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: Quantification of ELAVL4 mRNA using real-time quantitative RT-PCR. A: Standard curve used for ELAVL4 quantification ($R^2 = 0.999$; slope = -3.461). Ten-fold serial dilutions from NB cell line CLB-GA were evaluated (relative transcript number from 100,000 to 10). Each sample was analyzed in duplicate. The Ct values and relative transcript numbers were plotted on a logarithmic scale. (Ct: threshold cycle: cycle at which the fluorescence crosses the threshold) B: Amplification plot. ($\Delta$Rn: normalized fluorescence signal minus baseline fluorescence; RTN: relative transcript number)
Normalization:
Accurate normalization of gene expression levels is an absolute prerequisite to obtain reliable results. Because the expression of a single housekeeping gene can vary considerably, several references genes should be combined to calculate a normalization factor (33). In this study, TH, GALGT and ELAVL4 transcript numbers were normalized to two reference genes (beta-2-microglobulin (B2M) (34) and ubiquitin C (UBC)) by dividing the relative transcript number of the target genes by the geometric mean of the transcript number of the two reference genes.

PCR conditions:
All PCR reactions were performed on an ABI Prism 7700 Sequence Detection System (Applied Biosystems). A 3-by-3-primer matrix (combination of 300, 600 and 900 nM of each forward and reverse primer) was analyzed to determine the optimal primer concentrations. In addition, three different MgCl₂ concentrations (3, 5 and 7 mM) were evaluated. The concentrations resulting in the lowest threshold cycle and the highest fluorescent signal were chosen. PCR samples were prepared as follows. Five µl of cDNA was added to the reaction mixture and every sample was analyzed in duplicate. The amplification mixture for TH (25 µl) contained 1x TaqMan Buffer A (Applied Biosystems), 5 mM MgCl₂ (Applied Biosystems), 1.2 mM of each deoxynucleotide triphosphate (Amersham Biosciences), 150 nM of the probe, 600 nM of the forward and the reverse primer and 0.6 units of AmpliTaq Gold (Applied Biosystems). The amplification mixture for GALGT and ELAVL4 (25 µl) contained 1x TaqMan Buffer A, 7 mM MgCl₂, 1.2 mM of each deoxynucleotide triphosphosphate, 150 nM of the probe, 600 nM of both primer and 0.6 units of AmpliTaq Gold.

Due to their high expression, B2M and UBC were more efficiently amplified in 50 µl than in 25 µl. A reaction mixture (50 µl) containing 1x TaqMan Buffer A, 5 mM MgCl₂, 1.2 mM of each deoxynucleotide triphosphosphate, 150 nM of the probe, 600 nM of both primer and 1.25 units of AmpliTaq Gold was used.

The conditions for the PCR were 10 minutes at 95 °C, 50 cycles at 95 °C for 30 seconds and 60 °C for 1 minute.

Immunocytochemistry
Cytospins were immunocytochemically stained as described previously (16). Briefly, large diameter (17 mm) cytospins containing 5.10⁵ mononuclear cells were incubated with an unlabeled monoclonal antibody against GD2 disialoganglioside. GD2 is expressed by neuroectodermally derived tumors such as melanoma, neuroblastoma and small cell carcinoma but is not found in normal PB or BM cells. Controls where the primary antibody was replaced by buffer were included to evaluate the background staining. During a second and a third incubation step, cells were incubated with a polyclonal goat anti-mouse antibody
and a polyclonal rabbit anti-goat antibody, respectively. Both were labeled with alkaline phosphatase. The bound alkaline phosphatase complexes were stained using the new fuchsin chromogen. The cells were counterstained with hematoxylin and examined under a light microscope by two independent observers. Preferably six slides, each containing $5.10^5$ cells, were analyzed. Cells with a round nucleus, a granular chromatin structure, a scarce amount of cytoplasm and a strong staining around the entire cell were considered positive. Cells with an aberrant morphology or a pink staining restricted to a subcellular compartment were called negative. The number of positive single cells and positive cells being part of a Homer Wright rosette or cell clump were counted. The number of evaluated mononuclear cells was also reported. When the two local observers disagreed or less than 10 positive cells were found, pictures of the doubtful cell(s) were sent to an external expert. The results were discussed until agreement was reached. We were able to detect one NB cell in $10^6$ mononuclear cells using this immunocytochemical staining method (16).

**Statistics**

After the classification of the results in a frequency table, a $\chi^2$ test with Yates' correction for continuity was performed to study the relationship between the different categories. The Fisher's exact test was computed when a cell in the $2\times2$ table had an expected frequency less than 5. In addition, a Kappa test was used to measure the agreement between the different assays.
### RESULTS

**Expression of molecular markers in negative control samples**

The TH, GALGT and ELAVL4 expression was evaluated in 15 PB and 15 BM samples from patients without malignant disease. The results are summarized in Table 3.

Since TH, GALGT and ELAVL4 transcripts were detected in normal PB and BM samples, a threshold for positivity, based on the mean normalized transcript number and the standard deviation (SD), was defined. All clinical samples with a normalized relative transcript number higher than (mean + 4 SD) were considered positive.

**Table 3**: Normalized expression of TH, GALGT and ELAVL4 in 30 PB or BM samples taken from patients without malignant disease. All samples with a normalized relative transcript number higher than (mean + 4 SD) were considered positive. (PB: peripheral blood; BM: bone marrow; SD: standard deviation)

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Sample Type</th>
<th>TH</th>
<th>GALGT</th>
<th>ELAVL4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PB</td>
<td>0.00E+00</td>
<td>2.43E-05</td>
<td>2.52E-06</td>
</tr>
<tr>
<td>2</td>
<td>PB</td>
<td>0.00E+00</td>
<td>2.80E-05</td>
<td>7.74E-06</td>
</tr>
<tr>
<td>3</td>
<td>PB</td>
<td>0.00E+00</td>
<td>1.94E-05</td>
<td>1.45E-05</td>
</tr>
<tr>
<td>4</td>
<td>PB</td>
<td>2.24E-06</td>
<td>8.30E-06</td>
<td>1.81E-05</td>
</tr>
<tr>
<td>5</td>
<td>PB</td>
<td>0.00E+00</td>
<td>9.47E-06</td>
<td>2.78E-06</td>
</tr>
<tr>
<td>6</td>
<td>PB</td>
<td>0.00E+00</td>
<td>2.90E-05</td>
<td>9.42E-06</td>
</tr>
<tr>
<td>7</td>
<td>PB</td>
<td>0.00E+00</td>
<td>1.86E-05</td>
<td>1.51E-05</td>
</tr>
<tr>
<td>8</td>
<td>PB</td>
<td>0.00E+00</td>
<td>3.97E-05</td>
<td>1.56E-05</td>
</tr>
<tr>
<td>9</td>
<td>PB</td>
<td>0.00E+00</td>
<td>1.16E-05</td>
<td>1.26E-05</td>
</tr>
<tr>
<td>10</td>
<td>PB</td>
<td>0.00E+00</td>
<td>2.78E-05</td>
<td>9.41E-06</td>
</tr>
<tr>
<td>11</td>
<td>PB</td>
<td>0.00E+00</td>
<td>4.43E-05</td>
<td>5.36E-06</td>
</tr>
<tr>
<td>12</td>
<td>PB</td>
<td>0.00E+00</td>
<td>1.99E-05</td>
<td>7.92E-06</td>
</tr>
<tr>
<td>13</td>
<td>PB</td>
<td>0.00E+00</td>
<td>3.31E-05</td>
<td>4.93E-06</td>
</tr>
<tr>
<td>14</td>
<td>PB</td>
<td>0.00E+00</td>
<td>2.45E-05</td>
<td>8.37E-06</td>
</tr>
<tr>
<td>15</td>
<td>PB</td>
<td>0.00E+00</td>
<td>3.59E-05</td>
<td>2.59E-05</td>
</tr>
<tr>
<td>16</td>
<td>BM</td>
<td>1.80E-06</td>
<td>5.87E-05</td>
<td>1.77E-07</td>
</tr>
<tr>
<td>17</td>
<td>BM</td>
<td>0.00E+00</td>
<td>0.00E+00</td>
<td>0.00E+00</td>
</tr>
<tr>
<td>18</td>
<td>BM</td>
<td>0.00E+00</td>
<td>8.29E-05</td>
<td>6.38E-06</td>
</tr>
<tr>
<td>19</td>
<td>BM</td>
<td>0.00E+00</td>
<td>6.66E-04</td>
<td>3.15E-05</td>
</tr>
<tr>
<td>20</td>
<td>BM</td>
<td>0.00E+00</td>
<td>4.93E-04</td>
<td>1.66E-06</td>
</tr>
<tr>
<td>21</td>
<td>BM</td>
<td>0.00E+00</td>
<td>1.80E-03</td>
<td>0.00E+00</td>
</tr>
<tr>
<td>22</td>
<td>BM</td>
<td>0.00E+00</td>
<td>2.93E-05</td>
<td>0.00E+00</td>
</tr>
<tr>
<td>23</td>
<td>BM</td>
<td>0.00E+00</td>
<td>5.52E-06</td>
<td>1.55E-06</td>
</tr>
<tr>
<td>24</td>
<td>BM</td>
<td>0.00E+00</td>
<td>1.31E-04</td>
<td>0.00E+00</td>
</tr>
<tr>
<td>25</td>
<td>BM</td>
<td>0.00E+00</td>
<td>2.55E-04</td>
<td>1.73E-06</td>
</tr>
<tr>
<td>26</td>
<td>BM</td>
<td>0.00E+00</td>
<td>4.66E-05</td>
<td>0.00E+00</td>
</tr>
<tr>
<td>27</td>
<td>BM</td>
<td>0.00E+00</td>
<td>6.24E-05</td>
<td>2.17E-06</td>
</tr>
<tr>
<td>28</td>
<td>BM</td>
<td>0.00E+00</td>
<td>2.75E-04</td>
<td>1.37E-06</td>
</tr>
<tr>
<td>29</td>
<td>BM</td>
<td>0.00E+00</td>
<td>3.37E-05</td>
<td>0.00E+00</td>
</tr>
<tr>
<td>30</td>
<td>BM</td>
<td>0.00E+00</td>
<td>5.09E-05</td>
<td>7.41E-07</td>
</tr>
</tbody>
</table>

Mean 1.35E-07 1.45E-04 6.92E-06

SD 5.15E-07 3.46E-04 8.08E-06

Mean + 4 SD 2.20E-06 1.53E-03 3.92E-05
Sensitivity of the QPCR assays

The sensitivity of the TH, GALGT and ELAVL4 real-time quantitative RT-PCR assays was assessed by means of spiking experiments and clinical samples. The results are summarized in Table 4.

Different amounts of IMR32 RNA were added to 1 µg of RNA isolated from normal PB mononuclear cells. Five 10-fold dilutions (10,000 to 10 pg of IMR32 RNA) were evaluated. It was possible to detect 10 pg of IMR32 RNA in 1µg of normal RNA by means of TH and ELAVL4 QPCR. The sensitivity of the GALGT assay was lower. Only 100 pg of IMR32 RNA was reliably detected.

Both QPCR and immunocytochemistry were used to screen for residual neuroblastoma cells in patient samples. TH and ELAVL4 transcripts were detected in clinical samples containing only one neuroblastoma cell per 10^6 mononuclear cells. The sensitivity of the GALGT QPCR assay was lower. No GALGT transcripts were found in clinical samples containing less than 100 NB cells per 10^6 mononuclear cells.

Table 4: Sensitivity of the TH, GALGT and ELAVL4 QPCR assays. The sensitivity was evaluated using spiking experiments and clinical samples. (MNC: mononuclear cells, NB: neuroblastoma)

<table>
<thead>
<tr>
<th>Spiking Experiments (pg IMR32 RNA / µg normal MNC RNA)</th>
<th>Detection Limit Clinical Samples (# NB cells per million normal MNC's)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TH 10</td>
<td>1</td>
</tr>
<tr>
<td>GALGT 100</td>
<td>100</td>
</tr>
<tr>
<td>ELAVL4 10</td>
<td>1</td>
</tr>
</tbody>
</table>

TH, GALGT and ELAVL4 expression in clinical samples

The TH, GALGT and ELAVL4 expression was evaluated in 18 tumor, 70 BM, 5 PB and 4 PBSC samples from a GNB, a GN and 28 NB patients.

The three markers were highly expressed in all primary tumors (1 stage 1, 1 stage 2, 6 stage 3, 7 stage 4, 1 stage 4S and 1 GNB). Even two tumor samples from NB patients without elevated catecholamine levels scored positive for TH.

In addition, 32% (25/79), 11% (9/79) and 38% (30/79) of all BM, PB and PBSC samples scored positive for TH, GALGT or ELAVL4, respectively. Based on these results, a clear association between the different QPCR assays was found (TH and GALGT: Fisher’s exact test: p < 0.0001, κ = 0.44; TH and ELAVL4: \( \chi^2 = 20.15 \) and p < 0.0001, κ = 0.53; GALGT and ELAVL4: Fisher’s exact test: p < 0.0001, κ = 0.35).
**Comparison of real-time quantitative RT-PCR with immunocytochemistry**

The QPCR results of 68 BM, PB and PBSC samples were compared to those of an anti-GD2 immunocytochemical assay (IC). The latter was considered to be the gold standard.

When comparing TH or ELAVL4 QPCR with immunocytochemistry, only a few discordant results were found. There was a good association between the QPCR and immunocytochemical tests (TH and IC: $\chi^2 = 33.74$ and $p < 0.0001$, $\kappa = 0.74$; ELAVL4 and IC: $\chi^2 = 37.27$ and $p < 0.0001$, $\kappa = 0.77$) (Table 5).

More discordant results were found when GALGT QPCR and immunocytochemistry were compared. Fourteen samples were QPCR$^{-}$/IC$^+$, none were QPCR$^+$/IC$^-$. This reflects the difference in sensitivity between both assays. The association between the GALGT QPCR assay and the IC test was moderate (Fisher’s exact test: $p < 0.0001$, $\kappa = 0.46$).

Fourteen diagnostic BM samples from 9 stage 4 patients were analyzed. The results are summarized in Table 6. Seven samples scored positive for all three markers. The NB cells were also detected by IC and conventional cytomorphology (CM). In one BM sample, only TH and ELAVL4 were highly expressed. Another sample scored only positive for TH QPCR. In these samples no disseminated NB cells were found by CM or IC. Five BM samples were negative for QPCR, IC and CM.

The TH, GALGT and ELAVL4 expression was also evaluated in 12 diagnostic BM samples from 7 stage 3 patients. None of them scored positive for all three markers. However, in 4 samples ELAVL4 was highly expressed. One sample scored positive for TH QPCR. All samples were negative for IC and CM.

**Table 5:** Comparison of molecular and immunocytochemical results. (IC: immunocytochemistry)

<table>
<thead>
<tr>
<th></th>
<th>IC</th>
<th>TH</th>
<th>GALGT</th>
<th>ELAVL4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>19</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>4</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23</td>
<td>28</td>
<td>23</td>
</tr>
</tbody>
</table>

\[\chi^2 = 33.74\] \hspace{1cm} p < 0.0001

\[\kappa = 0.74\] \hspace{1cm} Good association

Fisher’s exact test : $p < 0.0001$, $\kappa = 0.46$ Moderate association

\[\chi^2 = 37.27\] \hspace{1cm} p < 0.0001

\[\kappa = 0.77\] \hspace{1cm} Good association
Potential prognostic value of detection of disseminated NB cells using real-time quantitative RT-PCR

To investigate whether the detection of TH, GALGT and ELAVL4 transcripts could be used to study MRD in NB, 30 subsequent BM, PB or PBSC samples from 3 stage 4 patients were studied. The samples were taken at diagnosis, during treatment, at relapse or during follow-up. The QPCR results were compared to immunocytochemical and cytomorphological data (Table 6).

In 14 out of 30 samples (47%) concordant results were found. Four samples scored positive in each assay. In 10 samples, no residual NB cells were found.

Patient 27 with subsequent positive results by IC, TH and ELAVL4 QPCR relapsed 14 months after diagnosis and died shortly thereafter.

In several PB and BM samples from patient 26 residual NB cells were found by IC, TH and ELAVL4 QPCR. This patient is still in complete remission at the conclusion of this study. However, follow-up time is short (3 months).

Patient 19 tested repeatedly negative at different sampling times and is still alive and disease free at the end of his treatment.

Table 6: Detection of TH, GALGT and ELAVL4 transcripts in BM, PB and PBSC samples from stage 3 and stage 4 patients. The samples were taken at diagnosis, during treatment, at relapse or during follow up. (Dx: diagnosis; CM: cytomorphology; IC: immunocytochemistry; BM: bone marrow; PB: peripheral blood; PBSC: peripheral blood stem cell preparation; NA: not analyzed; NE: not evaluable)
<table>
<thead>
<tr>
<th>Patient</th>
<th>Sample</th>
<th>Stage</th>
<th>Timepoint</th>
<th>Months after Dx</th>
<th>CM</th>
<th>IC</th>
<th>PCR TH</th>
<th>PCR ELAVL4</th>
<th>PCR GD2S</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>BM left</td>
<td>3</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM right</td>
<td>3</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>BM left</td>
<td>3</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM right</td>
<td>3</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>BM</td>
<td>3</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>BM</td>
<td>3</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>NE</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>BM left</td>
<td>3</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>NE</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM right</td>
<td>3</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>NE</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>BM left</td>
<td>3</td>
<td>Relapse</td>
<td>33</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM right</td>
<td>3</td>
<td>Relapse</td>
<td>33</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>BM left</td>
<td>3</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>NA</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM right</td>
<td>3</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>NA</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>BM left</td>
<td>3</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>NE</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM right</td>
<td>3</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>NE</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>BM left</td>
<td>4</td>
<td>On treatment</td>
<td>8</td>
<td>-</td>
<td>NA</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM right</td>
<td>4</td>
<td>On treatment</td>
<td>8</td>
<td>-</td>
<td>NA</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>BM left</td>
<td>4</td>
<td>Diagnosis</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM right</td>
<td>4</td>
<td>Diagnosis</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>BM left</td>
<td>4</td>
<td>Diagnosis</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM right</td>
<td>4</td>
<td>Diagnosis</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>BM left</td>
<td>4</td>
<td>Diagnosis</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM right</td>
<td>4</td>
<td>Diagnosis</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>BM</td>
<td>4</td>
<td>Follow-up</td>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>BM</td>
<td>4</td>
<td>Follow-up</td>
<td>11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>BM left</td>
<td>4</td>
<td>On treatment</td>
<td>2</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM right</td>
<td>4</td>
<td>On treatment</td>
<td>3</td>
<td>-</td>
<td>NE</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>BM</td>
<td>4</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>BM left</td>
<td>4</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM right</td>
<td>4</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>23</td>
<td>BM left</td>
<td>4</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM right</td>
<td>4</td>
<td>Diagnosis</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>BM left</td>
<td>4</td>
<td>On treatment</td>
<td>2</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>4</td>
<td>On treatment</td>
<td>3</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>25</td>
<td>BM</td>
<td>4</td>
<td>On treatment</td>
<td>3</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>26</td>
<td>BM</td>
<td>4</td>
<td>Diagnosis</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>4</td>
<td>On treatment</td>
<td>2</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>27</td>
<td>BM left</td>
<td>4</td>
<td>Diagnosis</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>BM right</td>
<td>4</td>
<td>Diagnosis</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>
DISCUSSION

The detection of residual NB cells in BM has important therapeutic and prognostic implications since BM disease is usually associated with unfavorable outcome (35, 36). Moreover, the detection of disseminated tumor cells in autologous stem cell preparations may signify the need for ex vivo manipulations such as tumor cell purging and/or CD34+ cell selection (37).

The cytomorphological screening of BM smears is still the only accepted technique for the detection of disseminated NB cells. However, the sensitivity of this method is limited. Therefore, alternative approaches, based on immunology or molecular biology, have been developed to improve the detection of rare NB cells.

In this study, disseminated NB cells were detected by means of TH, GALGT and ELAVL4 real-time quantitative RT-PCR. This technique offers several advantages compared to other methods. In contrast to conventional RT-PCR, no labor-intensive post-PCR processing such as gel electrophoresis and hazardous radioactive hybridization is needed to detect and quantify PCR products. In addition, carry-over contamination is minimized since both amplification and detection are performed within a closed system. Most importantly, real-time quantitative RT-PCR allows the detection of a very small number of transcripts and offers a wide dynamic range of quantification. As a consequence, this technique is particularly suited for monitoring changes in the tumor cell burden throughout the disease course.

We assessed the TH, GALGT and ELAVL4 expression in 18 primary NB tumors. Each tumor sample scored positive for TH, GALGT and ELAVL4. These findings suggest a possible applicability for TH, GALGT and ELAVL4 QPCR in the study of MRD in NB.

In tumor samples from 2 NB patients (stage 3 and 4S) with normal urinary catecholamine levels, a strong TH expression was found. This is in accordance with previously published results (38). The expression of the molecular markers was also evaluated in 79 BM, PB or PBSC samples. Respectively 32%, 11% and 38% of these samples scored positive for TH, GALGT or ELAVL4.

BM samples from 3 patients with a stage 3 neuroblastoma without MYCN amplification, scored positive for TH or ELAVL4 at diagnosis. In one of these patients a distal pathological lymph node was detected after diagnosis suggesting that this patient was stage 4 in stead of stage 3. Fourteen months after diagnosis, an additional BM sample from this patient was analyzed but no residual NB cells were found. All three patients are still alive and in complete remission at the conclusion of this study.

Using spiking experiments, we demonstrated that 10 pg of IMR32 RNA in 1µg of normal PB mononuclear cell RNA could be detected by TH or ELAVL4 QPCR. The sensitivity of the GALGT assay was about ten times lower.
The sensitivity of the molecular assays was also assessed using patient samples. QPCR results were compared to those of a standardized immunocytochemical assay which was considered to be the gold standard. We were able to detect TH and ELAVL4 transcripts in clinical samples containing only one NB cell per $10^6$ mononuclear cells. This is in accordance with previously published data (38). The sensitivity of GALGT QPCR was lower since no GALGT transcripts were found in clinical samples containing up to 100 NB cells per $10^6$ mononuclear cells. These results are in contrast with those published by Cheung and coworkers (39, 40). They evaluated the efficacy of adjuvant immunotherapy in NB patients using GALGT real-time quantitative RT-PCR and were able to detect one NB cell in $10^6$ normal mononuclear cells.

When TH or ELAVL4 QPCR was compared to immunocytochemistry, only a few discordant results were found. The discrepancies were probably due to sample variability (related to Poisson’s law) since only a few TH or ELAVL4 transcripts were detected in the QPCR+/IC- samples. In addition, less than 10 NB cells per $10^6$ mononuclear cells were found in QPCR-/IC+ samples.

More discordant results were found when GALGT QPCR and immunocytochemistry were compared. Fourteen samples scored positive for the IC assays. However, no GALGT transcripts were detected. This reflects the difference in sensitivity between both assays (IC $1.10^{-6}$ and GALGT $1.10^{-4}$).

The potential prognostic value of the detection of TH, GALGT and ELAVL4 transcripts was assessed analyzing 30 subsequent BM, PB and PBSC samples from 3 stage 4 NB patients. A 7.5 year old girl suffering from a stage 4 neuroblastoma with MYCN amplification and 1p deletion, tested repeatedly positive for IC, TH and ELAVL4 QPCR. In fact, ELAVL4 was highly expressed in every PB or BM sample taken during therapy. This child relapsed 14 months after diagnosis and died of malignant disease shortly thereafter. These preliminary data suggest that the persistence of a high ELAVL4 expression has prognostic value. Additional studies with larger patient groups are required to confirm these findings.

A 6 months old girl presented with a stage 4 neuroblastoma without MYCN amplification or 1p deletion. Several samples, collected at different time points during treatment, scored positive for TH and/or ELAVL4. This patient is still in complete remission at the conclusion of the study. However, follow-up time is limited (3 months).

A 6 year old boy diagnosed with a stage 4 neuroblastoma with 1p deletion and no MYCN amplification, tested repeatedly negative at different time points during therapy. This boy is still alive and in complete remission.

These data indicate that ELAVL4 and TH QPCR can be used to study bone marrow involvement in clinical samples collected during therapy or follow up. By contrast,
conventional cytomorphology and GALGT QPCR are not sensitive enough to study MRD and can only be used at diagnosis or relapse.

In conclusion, real-time quantitative RT-PCR for ELAVL4 mRNA can be used to study minimal residual disease in a sensitive and reliable way. However, we believe it is essential to combine multiple molecular markers (e.g. ELAVL4 and TH) with immunocytochemistry in order to avoid false positive or negative results.

ACKNOWLEDGEMENTS

The authors would like to thank Prof. Dr. F. Speleman for the provision of the neuroblastoma cell lines CLB-GA and IMR32. In addition, Prof. Dr. B. Verhasselt is gratefully acknowledged for his technical assistance and the use of the ABI PRISM 7700 Sequence Detection System. We are grateful to D. Claeys for his skilful assistance in the immunocytochemical assays.

Finally, the authors would like to thank Dr. B. Brichard, Dr. C. De Valck, Dr. N. Francotte, Dr. E. Michiels, Dr. A. Uyttebroeck and Dr. J. van der Werff for providing us with samples of neuroblastoma patients.
REFERENCES


2.4. Paper 7

DETECTING MINIMAL RESIDUAL DISEASE IN NEUROBLASTOMA PATIENTS: THE PRESENT STATE OF THE ART

Beiske K., Ambros P.F., Burchill S.A., Cheung I.Y. and Swerts K.

Cancer Letters (In Press)
Detecting minimal residual disease in neuroblastoma patients: 
the present state of the art

Beiske K., Ambros P.F., Burchill S.A., Cheung I.Y. and Swerts K.

Cancer Letters (In Press)

ABSTRACT

While cyto- and histological screening of bone marrow samples are still accepted as the gold standard for initial staging of neuroblastoma patients, these applications are insufficient during or after therapy because it is not always possible to detect tumour cell infiltration below the level of 1% by morphology alone. For monitoring of minimal residual disease, techniques offering a considerably higher sensitivity have been developed. Immunocytology, RT-PCR and flow cytometry are most frequently used, but differ with regard to targets (single cells, RNA transcripts), measured parameters (tumour cell number, antigen expression, cytomorphology, cytogenetic aberrations, level/number of RNA transcripts), specificity (uni-/multi-parameter analysis) and sensitivity (number of investigated cells). The pros and cons of these methods are reviewed. Precise quantification of residual tumour cells in bone marrow and blood may show a future impact on risk grouping and therapeutic strategies for patients with disseminated disease, but the potential clinical application of these techniques has to be preceded by thorough standardisation and validation in multi-centre studies.
Minimal Residual Disease

INTRODUCTION

Since bone marrow infiltration is frequently seen in neuroblastoma patients suffering from widespread disease, the detection of tumour cells at this site is a prerequisite for correct staging and monitoring of response to therapy [1,2]. For more than 20 years, morphological screening of bone marrow aspirates [3] and trephine biopsies [4] has been the gold standard, based on the identification of clumps or rosettes of tumour cells [5] or accompanying stromal changes [6]. Although the majority of comparative studies conceded a higher sensitivity to trephine biopsies, in particular when post-chemotherapy specimens were concerned, the representativity of biopsies was found to be inferior [7], with a high proportion of inadequate samples especially in infants [8]. Considering a detection level of 1% tumour cells for cytological and/or histological screening of bone marrow samples [9], it seems conceivable that morphological techniques alone were not regarded as reliable and sensitive enough to monitor minimal residual disease (MRD). For this purpose, an increasing number of new methods have been developed through the last two decades, based on the expression of biological, (i.e. immunological, cytogenetic and molecular genetic) markers by circulating tumour cells. Most of these techniques have shown a considerable higher sensitivity than morphology, and they have been applied to detect tumour cells not only in bone marrow, but also in peripheral blood. Provided that adequate controls are included, the identification of tumour cells by biological markers may also be expected to introduce a greater level of objectivity than morphological assays alone, and thus create the basis for a reliable quantification of single tumour cells. At the present, immunocytology with or without cytogenetic investigations, molecular genetic analyses (RT-PCR), and flow cytometry are the most frequently applied methods and will be briefly reviewed.

IDENTIFICATION OF CIRCULATING NEUROBLASTOMA CELLS BY IMMUNOCYTOMETRY

Immunocytological methods exploit the properties of monoclonal antibodies to bind to cell lineage- or differentiation-associated antigens expressed at high levels by tumour cells. Bound antibodies can be visualized by light or fluorescence microscopic detection methods and enable the morphological evaluation and quantification at the single cell level which is an essential prerequisite when the prognostic significance of very few residual tumour cells is addressed. However, due to the restricted specificity of some markers and to certain other technical aspects, e.g. shedding of the target antigen from tumour cells on to normal cells, not all immunocytological positive cells can be taken as representing tumour cells. Therefore,
Minimal Residual Disease

combinatorial techniques have been developed which aim to verify the neoplastic nature of immunostained cells either by co-expression of hematopoietic or non-hematopoietic markers or by the presence of cytogenetic abnormalities.

**Antigens**

Subsequent to the introduction of the hybridoma technique in 1978, monoclonal antibodies were raised against fetal brain and neuroblastoma cell lines [10,11,12,13], and were shown to recognize more or less well characterized antigens on neuroblastoma cells. In addition, mono- and polyclonal antibodies against established neuronal markers, e.g. neurone specific enolase [14], synaptophysin [15] and chromogranin A [16] were reported to bind to primary and metastatic tumour cells. However, regarding their reliability to detect MRD in bone marrow, many of these markers showed one or several restrictions, i.e. heterogeneous labelling of tumour cells in individual patients and cross-reactivity with normal hematopoietic cells [17] or differential expression on primary versus metastatic and on undifferentiated versus differentiated tumour cells [18]. To control cross-reactivity with hematopoietic cells, co-staining with anti-HLA-DR- [19], anti-CD10- [17] or anti-CD45- [20] antibodies were proposed. To overcome the heterogeneous inter- and intra-tumoural expression of individual markers, some researchers applied simultaneous staining with cocktails of antibodies of various specificities [21,22].

Among all markers summarized so far, two have gained special attention because of their relatively high sensitivity and/or specificity in MRD detection. Monoclonal antibody UJ13A was reported to be more sensitive and specific than other reagents produced by this group [23] and successfully applied also by others [20,24]. It recognizes the neural cell adhesion molecule CD56 (N-CAM) [25,26] which is expressed at very high levels by practically all neuroblastomas. In contrast to its high sensitivity, its specificity is restricted, because CD56 is also expressed by osteoblasts and – at considerable lower levels – by natural killer cells, subsets of plasma cells and macrophages [20, 27]. The other marker, disialoganglioside GD2, is a cell surface glycosphingolipid detected in the vast majority of neuroblastomas [13,28,29]. Undifferentiated neuroblasts are strongly positive, while differentiating tumour cells gradually lose the antigen [30]. GD2 is also expressed by melanomas, gliomas and focally in rhabdomyosarcomas and osteosarcomas [30]. GD2-specific antibodies were reported not to bind to normal hematopoietic cells [29] which render these reagents highly suitable for detection of MRD in the bone marrow. However, this antigen is not only down regulated during neuroblastic differentiation, but also continuously shed from the surface of neuroblasts [31]. Soluble antigen in the patients plasma can be ingested by bone marrow macrophages or unspecifically attach to the surface of hematopoietic cells and thus contribute to false positive results (Figure 1: D, E and F).
Antibodies against neuroblastoma-associated antigens have also been applied on sections of core biopsies in order to detect tumour cells by means of immunohistology. Some of these studies showed an increased sensitivity as compared to histomorphology alone [32, 33], but immunohistology never gained ground as a generally accepted technique for MRD detection. There may be at least three reasons why immunocytological analysis of aspirates prevailed: (1) the most specifically expressed molecule, GD2, is easily detected on unfixed, ethanol- or formaldehyde-fixed, but not on paraffin-embedded tumour cells. (2) As outlined in the introduction, biopsies are, in spite of being more sensitive, often less representative than aspirates, and the technical inadequacy of infant biopsies is well documented [8]. (3) The precise enumeration of a few positive cells and the obligatory assessment of nuclear/cytoplasmic details and subcellular localization of the immunolabel are easier achieved in cytological preparations of single cells than tissue sections. This brief review focuses therefore exclusively on cytological techniques.

**Figure 1:** A-F: Immunocytochemical detection of circulating neuroblastoma cells. Cytospin preparations of mononuclear bone marrow cells were stained with anti-GD2 antibody 14.G2a, alkaline phosphatase anti-alkaline phosphatase (APAAP) and Fuchsin + substrate. All images were taken with x60 objective and show identical final magnification. A, B: Criteria positive cells (CPC’s), i.e. tumour cells, frequently show a larger nucleus than lymphocytes, granular chromatin structure, scarce cytoplasm and strong GD2 membrane staining. C: This positive cell is considered as NCIC (not convincingly interpretable cell) because it is smaller than typical CPC’s and shows a weaker GD2 staining leaving out parts of the cell membrane (arrow). The true nature of this cell can only be disclosed by additional immunological or cytogenetic analyses. D: Two strongly GD2-positive CPC’s shed large amounts of antigen in between and on the surface of surrounding normal hematopoietic cells (arrows) simulating membrane staining of non-tumour cells. E, F: Shed GD2 antigen is taken up by macrophages and monocytes. Note striking cytoplasmic, but not membrane staining (arrows) of these cells which distinguishes them from CPC’s. G-H: Immunocytologic/cytogenetic analysis. This sample was stained with fluorescein-labelled anti-GD2 antibody (G) and subsequently hybridised with a fluorescein-labelled MYCN probe and a Rhodamin-labelled reference probe for chromosome 2 (H). Evaluation of each analytical step was performed by means of automated immunofluorescence plus FISH (AIPF). In spite of an only moderate and discontinuous GD2 membrane expression (G), the cell in the centre is proven to be a tumour cell by MYCN amplification (H).
**Immunological detection methods**

Both fluorescence- and light microscopy-based detection methods have been used [24, 34]. Each of them has its pros and cons. Light microscopic techniques usually involve enzyme/substrate reactions; alkaline phosphatase (AP) has been preferred to horseradish peroxidase due to the high amount of peroxidase in normal myelocytes. A light microscopic analysis enables the investigator to recognise established cytomorphological criteria for inclusion/exclusion of positively stained cells. Moreover, enzymatic staining results are stable over time and the immunocytological features of positively stained cells can be reviewed simultaneously by several investigators using a multi-headed microscope which is an essential element in the process of methodological standardisation between different laboratories. However, enzyme-based techniques are unsuited for the synchronous detection of a second antigen in the same subcellular localisation, i.e. a hematopoietic membrane antigen in addition to membrane GD2, because the molecular complexes of enzyme, substrate and chromogen cover an unpredictable number of neighboured epitopes like a shield. A combination with fluorochrome-labelled antibodies can neither be recommended due to the strong autofluorescence of most chromogens involved in AP/substrate reactions. Some individuals present in their bone marrow with very few normal plasma cells producing anti-AP antibodies [35] which could bind the enzyme/substrate complex in an unwanted manner. However, it is usually unproblematic to discriminate plasma cells from neuroblastoma cells on light microscopic grounds, and the presence of such false positive cells is detected in appropriate negative controls.

Fluorescence-based detection methods display features which by and large are inverse to those summarised for enzyme-based techniques. Fluorochromes are small molecules which do not require chemical reactions for visualisation. They are therefore ideal for the simultaneous demonstration of antigens, genes and chromosomes at the single cell level (see below). However, not all light microscopically defined cytomorphological features of neuroblastoma and hematopoietic cells can be recognised in a fluorescent microscope after nuclear staining with DNA-binding dyes e.g. DAPI. Moreover, the staining results cannot be reviewed in a multi-headed microscope and fade over time. They have to be documented by digital images which do not always contain all optical information about cell size and specific label as gained by direct fluorescence microscopy. Cheung and coworkers [36] reported four different staining patterns (background cytoplasmic, weak membrane, clear patchy membrane, and intense staining of the cell membrane), but no attempts have been published so far to standardise the criteria for evaluation and interpretation of fluorescent staining results for the purpose of MRD detection among different laboratories.
**Immunocytology and cytogenetics**

Ambros and Méhes took advantage of the properties of fluorescence-based systems to analyse simultaneously the expression of GD2 and of cytogenetic abnormalities in circulating neuroblastoma cells. They designed software for automated screening of immunofluorescent and fluorescent *in situ* hybridisation (FISH) signals in single cells [37]. The purpose of this approach was to prove the neoplastic origin of GD2–positive bone marrow cells by the simultaneous presence of e.g. hypersomy of chromosome 1 or 2, 17q gain, or MYCN amplification (fig. 1G and H). They showed that the very few (<10 per $10^5$) GD2–positive cells in bone marrow aspirates from stage 1, 2A and B, and stage 3 patients did not carry the cytogenetic aberrations known from the corresponding primary tumour, while those of all stage 4/4S patients did, and concluded therefore that the vast majority of GD2–positive cells in low stage patients do not represent neuroblastoma cells [38].

**Clinical significance**

Hallmarks of immunocytological assays for MRD detection are their high sensitivity and the enabling of a precise quantification of circulating tumour cells. Since even one single GD2-positive cell among $10^5$ or even more mononuclear cells can be demonstrated to be a neoplastic cell either by certain cytomorphological features, coexpression/lack of other antigens or cytogenetic aberrations, the sensitivity of this method depends on the number of investigated bone marrow cells and is therefore practically unlimited.

Some studies aimed to correlate the number of immunocytologically identified tumour cells to the patients’ outcome. Moss and co-workers [21] found a statistically significant impact on progression-free survival for stage 2 and 3 patients older than one year if their aspirates contained more than six tumour cells per $10^5$ bone marrow cells, while the corresponding cut-off for stage 4 patients below one year was 20 tumour cells per $10^5$. Seeger et al [22] performed the so far largest study on 466 patients with metastatic disease. Focusing on three-year event-free survival, they concluded that a cut-off of 100 tumour cells per $10^5$ bone marrow cells had a statistically significant predictive value when found in aspirates analysed before therapy, after 12 weeks, at bone marrow harvest and at the end of induction, but not after four weeks of treatment. Only at the time point of bone marrow collection, a lower number of tumour cells (20 per $10^5$) could herald an adverse prognosis. The lack of prognostic significance of small numbers of circulating tumour cells was ascribed to possible sampling variability due to differences in numbers of investigated bone marrow cells. In contrast to the data published by Moss et al, the number of bone marrow infiltrating tumour cells had no influence on the outcome of stage 3 patients.
Minimal Residual Disease

With regard to peripheral blood, Seeger et al confirmed previous reports indicating considerable lower numbers of circulating tumour cells as compared to bone marrow [39, 40]. The presence of more than one tumour cell per $10^6$ peripheral blood cells foreboded reduced event-free survival only when found at diagnosis, but not at the time of marrow collection.

In a recent study, Modritz et al [41] applied the combined immunocytologic/cytogenetic detection method in order to address the prognostic significance of bone marrow clearance in stage 4 patients at a detection level of one in $10^5$ mononuclear bone marrow cells. With regard to five-year survival, a statistically significant difference was found between patients obtaining bone marrow clearance within four cycles of chemotherapy and those who did not, even at the age of more than one year. However, the predictive value of rapid bone marrow clearance was only true for children with MYCN amplification and/or 1p deletion.

**Standardization of immunocytological detection**

The immunocytological studies reviewed above appear to disagree on a number of points, e.g. the presence and prognostic significance of tumour cells in the bone marrow of low stage patients, and the level of tumour cell infiltration which is able to predict the outcome of stage 4 patients before and during therapy. It seems highly probable that the differences in the results obtained by various laboratories are related to differences in their immunocytological methods, spanning from the number of investigated cells to details of the staining procedure and individual criteria for evaluation of staining results. In this respect, the Bone Marrow Subcommittee of the SIOP European Neuroblastoma Group has made a cooperative effort, based on contributions from laboratories in Austria, Belgium, France, Germany, Italy, Norway, Spain, Switzerland and UK, in order to standardise an immunocytological method for detection of MRD [42]. Since the sensitivity of immunocytological investigations varies with the number of investigated cells, the standard sample size was set to $3 \times 10^6$, thereby allowing a sensitivity of near one in $10^5$. Further standardisation steps comprised the same fixative (formaldehyde), a single monoclonal anti-GD2 antibody from the same commercial source, and the same detection method based on a commercially available alkaline phosphatase anti-alkaline phosphatase (APAAP) system. Since GD2 can be taken up or passively stick to normal bone marrow cells from neuroblastoma patients (fig. 1D, E and F), cytomorphological inclusion and exclusion criteria were developed. Positively stained cells were grouped into criteria-positive cells (CPC’s) (fig. 1A, B and D) or not convincingly interpretable cells (NCIC’s) (fig. 1C), a term applying to GD2-positive cells which neither fit unequivocally to CPC’s nor to any known hematopoietic category. The validity and reproducibility of this method was tested by quality control rounds where the participating laboratories analysed slides from the same aspirate and compared
their results. After introduction of the new standard method, the number of discordant results reported by individual laboratories decreased significantly \( p = 0.013 \). Another important experience gained from this work concerns the evaluation of samples with less than 10 GD2-positive cells per \( 10^6 \). Among so few stained cells, it can sometimes be difficult to identify cells which unequivocally display all inclusion criteria for CPC’s or form characteristic clumps. Therefore, these samples and those, which only contain NCIC’s, are subjected to an international central review performed by all collaborating laboratories at a multi-headed microscope. If no consensus is reached at this level, slides from such samples are analysed by the combined immunocytological/cytogenetic method which aims to clarify the neoplastic nature of positively stained cells by the presence or absence of cytogenetic aberrations. The subsequent steps of this standard procedure are summarised in the workflow scheme depicted in Figure 2. It is currently applied to investigate bone marrow aspirates and peripheral stem cell collections from patients enrolled in the European High Risk Study and it is expected to facilitate an unambiguous identification and quantification of tumour cells as a prerequisite for the assessment of prognosis and response to therapy in these patients.

**Figure 2:** Work flow for immunocytologic/cytogenetic MRD analysis within the European High Risk Study. Initial sample analysis is performed in the national laboratories applying the standardised immunocytochemical method as described in the text. Depending on the number of CPC and/or NCIC detected per \( 10^6 \) mononuclear cells (MNC), the stained slides are centrally reviewed and a decision is made whether cytogenetic analysis by AIPF (automated immunofluorescence plus FISH) must be performed.
**MOLECULAR DETECTION OF NEUROBLASTOMA**

Methods currently used in routine clinical practice for assessing disease status in neuroblastoma typically depend on radiographic studies and assessment of bone marrow for infiltrating neuroblastoma cells by Romanovsky stained smears and/or haematoxylin and eosin stained trephines. These methods are informative when extensive disease is present. However, for 10% of children with stage 4 disease at diagnosis, and most of them during follow up, these detection techniques may be suboptimal. Alternative methods with increased sensitivity to measure occult tumour cells are important for the early detection of metastases and relapse, and may also result in improved accuracy in staging and treatment strategies for some children with disease that is currently defined as low risk (stages 1-3). Such methods may identify the optimal timing for marrow or peripheral blood stem cell collection, and be useful tools to evaluate the efficacy of adjuvant therapies.

Numerous investigators have used reverse transcriptase polymerase chain reaction (RT-PCR) for tissue and tumour specific targets to detect minimal disease burden (reviewed in [43]). This is most specific when the target mRNA for RT-PCR is tumour specific e.g. the EWS-ETS gene rearrangement transcripts in tumours of the Ewing’s sarcoma family. However for most solid tumours, including neuroblastoma, such a tumour specific target has not yet been defined, leading to the exploitation of tissue specific targets. This approach requires the identification of target mRNA expressed in tumour cells but not in the compartments to be studied, namely bone marrow and peripheral blood. Using this methodology, clinically significant disease has been detected in peripheral blood [44, 45, 46, 47, 48] and bone marrow [22, 44, 46, 49, 50, 51, 52, 53] samples from children with neuroblastoma at diagnosis, on therapy, during follow up and at relapse. From such studies the persistent expression of target mRNA in bone marrow or peripheral blood from children with stage 4 disease despite intensive chemotherapy, is predictive of relapse [47, 53] and suggests the early use of novel therapeutic strategies may be appropriate in these children. Such an approach may also provide an early indicator of response and treatment efficacy, for example in patients receiving adjuvant immunotherapy [54].

Neuroblastoma cells express catecholamines, consequently the first enzyme in the catecholamine synthesis pathway tyrosine hydroxylase (TH) has successfully been exploited for the detection of clinically significant neuroblastoma cells by RT-PCR [55]. Other targets for RT-PCR, such as GD2 synthase, the key enzyme for the synthesis of GD2 (a ubiquitous and abundant antigen on neuroblastoma [51]), have shown promise as molecular markers. Although RT-PCR for TH mRNA detects clinically significant disease with increased sensitivity and specificity compared to more conventional methods, it is clear that tumour cells are not detected in bone marrow or peripheral blood of all children with high-risk
disease. This has been interpreted by some as a “false” negative result, although the absence of tumour cells in these samples may reflect the biology of tumour cell shedding, dilution of tumour cells within the compartment and technical features such as the volume and frequency of samples analysed. In some tumours this might also reflect differences in target mRNA expression, suggesting that the detection of tumour cells by RT-PCR may be enhanced by using multiple molecular markers to overcome any tumour heterogeneity [56]. It is therefore important to identify and evaluate the sensitivity and specificity of new markers, as well as to validate their clinical utility.

The development of semi-quantitative methods, comparing levels of target mRNA to those of a house-keeping gene such as β2-microglobulin, has allowed potentially useful evaluation of changes in tumour cell burden within peripheral blood of children throughout the disease course (Figure 3A) [47]. Recent studies have also used quantitative real-time RT-PCR to determine target gene transcript level in clinical samples from children with neuroblastoma (Figure 3B) [51, 57, 58, 59]. Such quantitation may allow more precise risk assessments in disease monitoring. Real-time quantitative PCR is an attractive technology, with a wide linear dynamic range, superior sensitivity, high throughput capacity, and elimination of post-PCR handling steps. The selection of an optimal reference gene against which the test samples can be normalised is essential [60]; the use of target specific calibrators that allow the accurate quantification of transcript number within a sample is desirable [57]. Nevertheless, the high sensitivity of the real-time PCR assay can lead to the detection of illegitimate target gene transcripts and an increased number of false positives with some targets. Thus, clinical validation of these markers detected by real-time RT-PCR is vital.

Although the capacity to detect neuroblastoma micrometastases does not by itself inevitably translate into improved outcome for children with this disease, the ability to detect occult tumour cells in the bone marrow and peripheral blood may allow evaluation of new strategies targeting this disease (Figure 3C) [48, 55]. The detection of clinically relevant micrometastases may also result in a redefinition of what constitutes remission and relapse, possibly leading to improved risk-stratification of children for a better curative therapy. However, the clinical utility of molecular detection still remains unclear, reflecting the relatively small number of children studied and the inconsistency in methodology and reporting. Thus, there is a need for large, multi-centre, quality controlled clinical outcome efficacy studies to maximize the clinical benefit of RT-PCR detection of disease for children with neuroblastoma.
Figure 3: A: Detection of TH mRNA by RT-PCR in sequential blood samples (1-10) from a child with stage 4 neuroblastoma. (RT+ve = reverse transcriptase present, RT-ve = reverse transcriptase absent (negative control), TH = tyrosine hydroxylase mRNA. β2m = β2 microglobulin mRNA (house-keeping gene), W=water negative control. C¹ = RNA from neuroblastoma cell line, C² = 10 neuroblastoma cells spiked into 2ml of whole blood (positive controls), M=molecular weight markers.) B: Amplification plot generated by real-time RT-PCR for GD2 synthase mRNA, and quantification (against standards) of GD2 synthase mRNA in bone marrow samples from children with stage 4 neuroblastoma. C. Potential clinical utility of RT-PCR detection of neuroblastoma cells.
DETECTION OF RESIDUAL NEUROBLASTOMA CELLS USING CYTOMETRY

The first triple-colour flow cytometric assay for the detection of disseminated neuroblastoma cells in peripheral blood or bone marrow was developed by Komada and coworkers [61]. The assay was based on the detection of CD9, CD56 and CD45.

CD9, a 24 kDa Type IV transmembrane protein belonging to the family of tetraspanins, is expressed on platelets, pre-B cells, activated and differentiating B cells, activated T cells, eosinophils, basophils, monocytes, endothelial and epithelial cells, brain, peripheral nerves and muscular cells. Various in vitro studies confirmed the consistent expression of CD9 on neuroblastoma cells [62]. The CD56 antibody reacts with 175/220 kDa glycosylated antigen isoforms present on natural killer cells. It is also present on a subset of CD4⁺ and CD8⁺ T cells in peripheral blood, neural derived cells and tumours, large granular lymphocyte and myeloid leukaemias, small cell lung carcinomas and myelomas. The neural cell adhesion molecule (NCAM) is also expressed on neuroblastoma cells [63]. The leukocyte common antigen CD45, present on all human leukocytes but absent on neuroblastoma cells, was included as a negative marker. Sorted CD9⁺/CD56⁺/CD45⁻ cells were considered to be true neuroblastoma cells since all of them showed a neuroblast-like morphology. Using this triple-colour flow cytometric assay, Komada was able to detect one neuroblastoma cell in 10⁴ to 10⁵ mononuclear cells.

Komada's flow cytometric assay was modified by Nagai et al. [64]. He replaced CD9 with CD81 because he suspected that the interaction between CD9 and platelets could lead to platelet aggregation. CD81, a 26 kDa member of a multimeric cell surface signal transduction complex, is expressed on hematopoietic, endothelial and epithelial cells. The expression of CD81 on neuroblastoma cells has also been described [65]. Based on comparative studies, the combination of CD81⁺/CD56⁺/CD45⁻ was found to be more sensitive and specific than CD9/CD56/CD45 for the detection of disseminated neuroblastoma cells. Nagai was able to detect CD81⁺/CD56⁺/CD45⁻ cells at a concentration of 0.005%.
Minimal Residual Disease

Warzynski and coworkers evaluated several combinations [66]. Initially, a cocktail of CD45, CD56 and cytoplasmic neuron specific enolase (NSE) was used to detect rare neuroblastoma cells. This procedure was later refined and modified to a less laborious anti-GD2/CD56/CD45 test eliminating the cell permeabilization step for cytoplasmic NSE analysis. GD2 disialoganglioside is highly expressed by neuroectodermally derived tumours such as melanoma, neuroblastoma, small cell carcinoma and glioma. GD2 expression is not found in normal bone marrow or peripheral blood samples [36]. Finally, CD81 was added to the previous combination. Using patient samples instead of artificially spiked material, Warzynski was able to detect neuroblastoma cells in bone marrow at a level of 0.002%.

Tsang et al. compared the triple-colour CD81/CD56/CD45 flow cytometric assay to tyrosine hydroxylase reverse transcriptase-polymerase chain reaction (RT-PCR) and found a significant correlation between both techniques [67]. Given the advantage of flow cytometry in analysing large cell numbers, Tsang recommended to combine flow cytometry with qualitative RT-PCR for TH to assess tumour regression and to monitor minimal residual disease in neuroblastoma patients.

A combination of two four-colour flow cytometric assays (CD9/CD81/CD56/CD45 and anti-GD2/CD81/CD56/CD45) was used to detect residual neuroblastoma cells in a study published by Swerts et al. [68]. Based on spiking experiments, a sensitivity of $1.10^{-5}$ was reached. When flow cytometric results were compared to immunocytochemical data, a strong correlation between both assays was found.

Flow cytometry has been used in the clinical practice of haematology and oncology for many years and has some advantages compared to immunocytochemistry or RT-PCR. Flow cytometry provides information on multiple cellular parameters such as cell size, granularity and marker expression in a multicolour assay. In addition, it allows the identification of dying cells and cellular debris which makes it possible to exclude irreversibly damaged cells which are unable to expand. Large numbers of cells can be rapidly screened and the technique is relatively simple and cost-effective. One specific advantage of flow cytometry over PCR based assays is that it allows direct quantification of MRD rather than extrapolating it from amounts of PCR products. This makes quantification easier and more accurate.

Flow cytometric assays have also limitations. In contrast to immunocytochemistry or RT-PCR, it is virtually impossible to detect one neuroblastoma cell in $10^5$ or more mononuclear cells (0.0001%). However, such an extreme sensitivity may not be relevant. Seeger and coworkers reported that the failure to reduce neuroblastoma cells in bone marrow to fewer than 100 tumor cells per $10^5$ bone marrow (0.1%) cells within 12 weeks after start of treatment has a prognostic value, indicating that the sensitivity of most flow cytometric assays is sufficient to study MRD in neuroblastoma patients [22].
CONCLUSIONS

Immunocytology, RT-PCR and flow cytometry are the most frequently applied techniques for detecting MRD in neuroblastoma patients. The sensitivity of cell counting methods (immunocytology and flow cytometry) is in the range of 1 tumour cell per $10^5$ to $10^6$ mononuclear cells or even higher and thus clearly beyond the sensitivity of cyto-/histomorphological screening. Immunocytology is the oldest of these techniques, and enables both a critical cytomorphological evaluation and precise quantification of positively stained cells. The simultaneous detection of additional immunological and cytogenetic markers is an important tool (1) to exclude false positive cells and (2) to analyse circulating tumour cells for other biological and possibly prognostic significant properties. The sensitivity of this method is practically unlimited. Also flow cytometry measures the expression of antigens on single cells and can easily be used to detect several markers at the same time in order to increase the specificity and/or biological information provided by the analysis. Compared to immunocytology, flow cytometry is able to screen a much larger number of cells in a shorter time, but its sensitivity is limited (1 in $10^5$). RT-PCR is a highly sensitive technique which can demonstrate the presence of low levels of tumour-specific RNA transcripts in blood and bone marrow. Semi-quantitative RT-PCR and quantitative real-time RT-PCR have been designed to analyse the relative level and even the exact number of tumour-specific transcripts, respectively. Up to now, the differences between the parameters measured by each technique make it difficult to directly compare the sensitivity of cytologic/flow cytometric analyses to the one of molecular genetic assays, mainly because of two reasons: (1) Due to a possible heterogeneous gene expression between and within various tumours, the amount of transcripts identified by RT-PCR does not necessarily correspond to the number of tumour cells in the sample. (2) Results of molecular genetic analyses are often reported in relation to the volume of investigated blood/bone marrow, but not to the total cell number in the sample which may vary tremendously depending on the phase of treatment. As outlined previously, the number of investigated cells is an essential parameter for calculating the sensitivity of cytologic and flow cytometric investigations. Despite all the technical progress which has been made to detect small numbers of circulating tumour cells, the clinical significance of these observations remains unclear. Published reports disagree on the numbers of bone marrow infiltrating tumour cells which appear to have prognostic significance. Large multi-centre studies are needed to address these issues and may be expected to identify cut-off levels of tumour cell infiltration/RNA transcripts which could be useful to define risk groups and to tailor therapeutic strategies for patients suffering from MRD. However, such co-operative efforts can only reach reliable conclusions if the involved laboratories apply techniques of identical sensitivity and
specificity. It is therefore crucial that all participants agree upon a standardisation of their method before starting on this work. The SIOP Europe Neuroblastoma Bone Marrow Subcommittee has recently proposed a standardisation of the immunocytological method designed for the analysis of bone marrow and peripheral stem cell harvests of children within the European High Risk Study, and similar efforts are in progress to standardise molecular genetic investigations.
REFERENCES


3. Clinical importance of neuroblastoma micrometastases

3.1. Staging and risk assessment at diagnosis

Evaluation of bone marrow aspirates and biopsies is essential for clinical staging and risk assessment at diagnosis. The presence of neuroblastoma cells in diagnostic bone marrow samples identifies those children with an unfavorable prognosis by association with disease stage. Standard cytology reveals bone marrow metastases in 80% to 90% of stage 4 patients. By contrast, 95% to 100% of these children are positive when analyzed by immunocytochemistry or RT-PCR suggesting that the prevalence of bone marrow involvement is underestimated by conventional cytomorphology. The detection of occult neuroblastoma cells in diagnostic bone marrow samples from children with stage 4 disease is unlikely to have an effect on outcome since these patients are already classified as high risk and treated accordingly. By contrast, it might be clinically relevant to use sensitive techniques to screen for bone marrow involvement in children with apparently localized disease. Whether or not these patients would benefit from upstaging and more aggressive treatment remains to be evaluated.

Several research groups evaluated the clinical importance of circulating neuroblastoma cells in peripheral blood at diagnosis. Seeger and colleagues found a significant association between circulating neuroblastoma cells and unfavorable outcome in stage 4 patients over 1 year of age. These results were confirmed by Burchill and coworkers. In addition, detection of circulating tumor cells in peripheral blood of stage 3 patients appears to correlate with poor prognosis.

3.2. Monitoring disease course and response to therapy

Several reports link the detection of minimal residual disease in bone marrow during therapy to outcome in high-risk neuroblastoma patients. Saarinen and coworkers used a sensitive immunologic assay to monitor in vivo bone marrow purging in children with high-risk neuroblastoma. The purging efficacy during the first 3 months after diagnosis was a strong prognostic factor reflecting tumor responsiveness to therapy. Seeger and colleagues evaluated bone marrow samples from 466 children with high-risk metastatic neuroblastoma and reported that patients with more than 100 tumor cells per 10^5 nucleated bone marrow cells after three to four cycles of chemotherapy had virtually no chance of survival. Fukuda et al. published similar results. The persistence of MRD in bone marrow, four months...
after the start of chemotherapy, was a poor prognostic parameter in high-risk neuroblastoma.\textsuperscript{203,209} Tchirkov and coworkers used tyrosine hydroxylase (TH) real-time RT-PCR to study residual disease in bone marrow. The overall survival for patients with more than 1000 TH copies in bone marrow after initial chemotherapy was significantly worse.\textsuperscript{210} Cheung and colleagues evaluated the clinical relevance of micrometastases in bone marrow at the end of treatment.\textsuperscript{211,212} Their studies showed that detection of GAGE or GD2 synthetase transcripts in bone marrow, 24 months after completion of intensive therapy combined with anti-GD2 antibody immunotherapy, correlated with poor event-free survival. The detection of residual neuroblastoma cells in peripheral blood from clinically disease-free children is also associated with rapidly progressing disease and high relapse rate.\textsuperscript{206} Preliminary studies have shown that neuroblastoma cells can be detected in peripheral blood up to 11 months before urinary catecholamine concentrations increase and clinical relapse occurs.\textsuperscript{206} Whether treatment of this low-level disease will offer a survival advantage remains unclear.

\section*{3.3. Screening of autologous stem cell preparations}

High-dose chemotherapy and subsequent hematopoietic rescue are widely used in the treatment of high-risk neuroblastoma patients. Despite this intensive therapy, many children relapse. The clinical significance of re-infusing neuroblastoma cells during transplantation procedures has been a matter of debate. However, recent studies have shown that circulating neuroblastoma cells are highly colonogenic and, if re-infused, could lead to system recurrence.\textsuperscript{199,201,213} In addition, Tchirkov and colleagues found a relationship between increased levels of TH transcripts in peripheral blood stem cell (PBSC) harvests and reduced survival.\textsuperscript{210} A similar tendency was observed by Burchill and coworkers.\textsuperscript{214} The correlation between tumor cell detection in PBSC harvests and poor post-transplant clinical outcome supports a role for re-infused tumor cells in relapse. Consequently, the use of purged PBSC collections is believed to reduce the risk of relapse post-transplant. However, it remains unclear whether such procedures would offer a clinical advantage.
4. Discussion and future perspectives

Conventional cytomorphology is still the only accepted technique for the detection of disseminated neuroblastoma cells in bone marrow, peripheral blood and PBSC harvests. This method is informative at diagnosis or relapse when overt metastases are present. However, it cannot be used to detect occult neuroblastoma cells in clinical samples collected during therapy or follow up because of its limited sensitivity (at the best $1.10^{-3}$). Consequently, sensitive methods for the detection of rare neuroblastoma cells and the evaluation of their clinical significance are urgently needed. Assays based on immunocytochemistry, flow cytometry and real-time quantitative RT-PCR were developed and evaluated (reviewed in Paper 7).

In collaboration with the SIOPEN Bone Marrow Subcommittee, we developed, optimized and standardized an immunocytochemical assay, formulated morphological and immunological criteria and designed a work flow (Paper 4). Four important improvements were achieved: 1) a sensitivity of $10^{-6}$ was reached, 2) the concordance between the different observers, with regard to staining and evaluation of results, increased significantly, 3) the range between the highest and lowest reported result was reduced by half and 4) discordant results were only found in samples with less than 10 CPC’s per $1.10^6$ mononuclear cells. We describe the first internationally standardized protocol for detection and quantification of occult neuroblastoma cells by immunocytochemistry. Our method is a valuable tool for multicenter studies evaluating the clinical significance of minimal residual disease in neuroblastoma.

In paper 5, two four-color flow cytometric assays using different combinations of CD9, CD81, CD56, CD45 and anti-GD2 disialoganglioside are described. The sensitivity of both tests was assessed using spiking experiments and the results were compared to those of an immunocytochemical reference assay. The flow cytometric assays were less sensitive than the immunocytochemical test which was considered to be the gold standard. However, as few as one neuroblastoma cell in $10^4$ to $10^5$ mononuclear cells could be detected. Since Seeger et al. reported that a sensitivity of 0.1% after three to four cycles of therapy most likely represents the clinically useful threshold, we still believe that flow cytometry can be used to study MRD in neuroblastoma.
Within the framework of this thesis, disseminated neuroblastoma cells were also detected by means of real-time quantitative RT-PCR (QPCR) (Paper 6). The expression of three different genes (tyrosine hydroxylase, GALGT and ELAVL4) was evaluated. We were able to detect ELAVL4 and TH transcripts in clinical samples containing only one neuroblastoma cell per $10^6$ mononuclear cells whereas no GALGT transcripts were found in samples containing up to 100 neuroblastoma cells per $10^6$ mononuclear cells. The potential prognostic value of TH, GALGT and ELAVL4 QPCR was assessed by analyzing subsequent samples from stage 4 patients. Preliminary results indicate that persistence of high ELAVL4 expression is associated with unfavorable outcome. However, further studies are needed to confirm these results.

Several research groups found a relationship between sensitive detection of disseminated neuroblastoma cells in peripheral blood, bone marrow and PBSC harvests during or after therapy and unfavorable outcome. On the basis of these data, one might be tempted to use the detection of residual neuroblastoma cells during or after treatment as a prognostic factor to identify patients at risk for relapse. However, prudence is called for since therapies used after MRD detection may decrease the significance of neuroblastoma micrometastases. Indeed, Seeger and coworkers reported that residual neuroblastoma cells detected in bone marrow harvested for autologous hematopoietic stem cell transplantation (AHSCT) in patients who were not treated with the differentiation inducer 13-cis retinoic acid (13-cis-RA) correlated with poor EFS. However, this was not the case for patients randomized to receive 13-cis-RA. Other differences in treatment regimens such as purging of stem cell harvests could also influence the clinical significance of MRD detected during induction therapy. By contrast, detection of tumor cells in peripheral blood or bone marrow after completion of therapy may be less influenced by differences in therapeutic approach. However, the prognostic value of neuroblastoma micrometastases in post-treatment samples from patients treated with 13-cis-RA could differ from the clinical relevance of residual tumor cells in samples from children not treated with this differentiation inducer. Viable but terminally differentiated non-proliferating tumor cells could persist in bone marrow or peripheral blood for months to years after therapy without leading to progressive disease.

Large and well controlled clinical studies are indispensable to determine the prognostic significance of disseminated neuroblastoma cells in bone marrow, peripheral blood or PBSC collections. MRD evaluation before, during and after therapy is an integral part of the ongoing high risk study of the International Society of Pediatric Oncology European Neuroblastoma (SIOPEN) group. Bone marrow, peripheral blood and PBSC samples, collected at various time points during therapy, will be analyzed by immunocytoology and RT-PCR. The large
number of uniformly treated patients will provide the statistical power needed to define the prognostic value of MRD in neuroblastoma. However, large cooperative studies can only reach reliable conclusions if all involved laboratories use standardized detection techniques.

In conclusion, accurate detection of disseminated tumor cells will improve staging, risk assessment, therapy stratification and monitoring of therapeutic response in neuroblastoma patients. Moreover, the more sensitive and specific detection of micrometastases may ultimately lead to a redefinition of important clinical terms such as ‘disease free’ and ‘relapse’. We believe large multicenter, quality controlled, prospective clinical studies are indispensable to determine the clinical usefulness of MRD detection in neuroblastoma. Such studies will generate data which will ultimately lead to an improved outcome for children.
References


References


References


References


Summary

Well designed multimodal treatment regimens have led to a significant improvement in long-term event-free survival of children diagnosed with cancer. However, a substantial number of patients relapses and cannot be cured with modern therapy. The identification of features predicting a patient’s response to therapy has been a continuing goal of researchers working in the field of pediatric hematology-oncology. These prognostic parameters are used for risk assessment and therapy stratification: low risk patients are treated with less intensive and toxic chemotherapy, whereas high risk patients receive a more aggressive treatment. Recently, two approaches have become popular: 1) screening for transmembrane protein-mediated multidrug resistance (MDR) in childhood ALL and 2) minimal residual disease (MRD) analysis in neuroblastoma.

Intrinsic or acquired resistance, mediated by transmembrane transport proteins, is one of the most important causes of treatment failure in childhood ALL. Several multidrug resistance-related drug efflux pumps have been characterized (e.g. P-glycoprotein, multidrug resistance-associated protein, lung resistance protein and breast cancer resistance protein) but their clinical relevance in childhood ALL remains controversial. The lack in standardized detection techniques hampers the comparison of data. Moreover, most assays are not sensitive enough to detect low-level resistant cells. Consequently, one of our major goals was the development, evaluation and optimization of a more sensitive and specific detection technique for the assessment of P-glycoprotein activity. In addition, the clinical relevance of P-glycoprotein expression and activity in childhood ALL was evaluated in a prospective study and current knowledge about the clinical relevance and prognostic value of multidrug resistance-related proteins in pediatric ALL patients was summarized. Several researchers found a relationship between P-gp, LRP, BCRP, and MRP3 expression and/or activity and an unfavorable prognosis in childhood ALL. However, larger multicenter studies are needed to confirm these results.

Reliable detection of neuroblastoma cells in bone marrow and peripheral blood is crucial for correct clinical staging and risk assessment at diagnosis and monitoring therapeutic response during treatment. Moreover, screening of autologous stem cell preparations is crucial since the re-infusion of contaminated stem cell products could lead to systemic recurrence. Conventional cytomorphological screening of bone marrow aspirates and biopsies is only informative at diagnosis or relapse when overt metastases are present but it cannot be used to detect occult neuroblastoma cells in clinical samples collected during
therapy or follow up because of its limited sensitivity. Consequently, sensitive methods for the detection of rare neuroblastoma cells and the evaluation of their clinical significance are urgently needed. As part of this study, sensitive and specific assays based on immunocytochemistry, flow cytometry and real-time quantitative RT-PCR were developed, evaluated and standardized. In addition, data on clinical importance of disseminated disease was reviewed. Several research groups found a relationship between sensitive detection of disseminated neuroblastoma cells in peripheral blood, bone marrow and PBSC harvests during or after therapy and unfavorable outcome. However, international, quality controlled, prospective clinical studies are indispensable to confirm these preliminary results.

In conclusion, the detection techniques developed within the framework of this thesis allow the accurate assessment of multidrug resistance and minimal residual disease in childhood malignancies. This will improve risk classification and therapy stratification and may ultimately result in increased long term survival for children with cancer. In addition, an overview of the clinical relevance of the above mentioned prognostic parameters is given.
Samenvatting

Door de ontwikkeling van efficiëntere behandelingsschema's is de lange termijn overleving van kinderen met kanker significant verbeterd. Toch volstaan de beschikbare therapeutische middelen niet om alle patiënten te genezen. Men is op zoek naar klinische, biologische en/of genetische prognostische parameters om kinderen met kanker in verschillende risicogroepen in te delen. Patiënten met goede prognostische eigenschappen worden minder intensief behandeld terwijl kinderen met ongunstige prognostische factoren agressievere therapie ontvangen. Recent werden twee nieuwe prognostische parameters beschreven. Enerzijds zou de expressie en activiteit van resistentie-gerelateerde transporteiwitten in leukemische cellen met een ongunstige prognose gepaard gaan. Anderzijds zou er een verband bestaan tussen de detectie van residuele neuroblastoomcellen in beenmerg (afgenomen tijdens de behandeling) en een verhoogde kans op herval.

Intrinsieke of verworven chemotherapeutieresistentie is één van de meest belangrijke oorzaken van therapeutisch falen bij kinderen met ALL. Er kunnen verschillende mechanismen aanleiding geven tot multidrug resistentie (MDR), maar cellulaire resistentie, veroorzaakt door overexpressie van transmembranaire transporteiwitten, is het best bestudeerd. Deze transporteiwitten pompen chemotherapeutica uit de cel en verminderen zo de intracellulaire toxiciteit. Er zijn verschillende MDR-pompen beschreven (o.a. P-gp, MRP1, BCRP en LRP) maar hun klinisch en prognostisch belang bij ALL-patiëntjes blijft onduidelijk ondanks diepgaand onderzoek. Door een gebrek aan gestandaardiseerde detectietechienen is het moeilijk om de resultaten van de verschillende studies met elkaar te vergelijken. Bovendien zijn de gebruikte testen vaak niet gevoelig genoeg om weinig resisteente tumorcellen op te sporen. Daarom werd er in het kader van deze thesis een nieuwe, gevoelige en specifieke techniek ontwikkeld om de P-gp activiteit te bepalen. Het klinische belang van P-gp expressie en activiteit bij kinderen met ALL werd in een prospectieve studie geëvalueerd en de bestaande informatie over het klinische en prognostische belang van de andere multidrug resistentie eiwitten werd samengevat. Verschillende onderzoekers vonden een verband tussen P-gp, LRP, BCRP en MRP3 expressie en/of activiteit en een ongunstige prognose bij kinderen met ALL. Verder onderzoek zal uitwijzen of deze resultaten bevestigd kunnen worden.

De betrouwbare detectie van tumorcellen in beenmerg van neuroblastoompatiënten is zeer belangrijk in het kader van een correcte stagering en risico-evaluatie op het moment van diagnose en de bepaling van de therapeutische respons tijdens de behandeling. Daarnaast
Samenvatting

is ook het screenen van autologe stamcelcollectes van belang omdat de reïnfusie van tumorcellen aanleiding kan geven tot herval. De cytomorfologische beoordeling van beenmerg uitstrijkjes is enkel informatief bij diagnose of herval wanneer het beenmerg massaal ingenomen is. Door een gebrek aan gevoeligheid kan deze techniek echter niet gebruikt worden om residuele tumorcellen op te sporen tijdens de therapie. Meer gevoelige en specifieke detectietechnieken zijn dan ook dringend nodig. Wij ontwikkelden en standaardiseerden nieuwe specifieke en gevoelige methodes op basis van immunocytochemie, flowcytometrie en ‘real-time’ kwantitatieve RT-PCR. Daarnaast werd het klinische belang van minimale residuele ziekte bij kinderen met neuroblastoom bestudeerd. Verschillende onderzoekers toonden aan dat residuele ziekte met een verhoogde kans op herval geassocieerd is. Internationale, gestandaardiseerde studies zijn echter nodig om deze preliminaire resultaten te bevestigen.

Met behulp van de detectietechnieken die in het kader van deze thesis ontwikkeld werden, kan het belang van multidrug resistentie en minimale residuele ziekte bij kinderen met kanker nauwkeuriger bepaald worden. Dit kan tot een verbeterde risico-evaluatie en therapiestratificatie leiden wat uiteindelijk moet resulteren in een verbeterde lange termijn overleving voor kinderen met kanker.
Résumé

Grâce au développement récent de schémas thérapeutiques combinés, la survie à long terme des enfants atteints de cancer s’est considérablement améliorée. Cependant, même les thérapies actuelles ne permettent pas encore la guérison d’un nombre significatif de patients en rechute. La recherche de facteurs biologiques et génétiques prédictifs d’une réponse au traitement représente un objectif majeur pour les pédiatres oncologues. L’identification de marqueurs à haute valeur pronostique permet d’orienter le traitement et de classer plus précisément les patients dans différents groupes de pronostic. Ainsi un patient appartenant à un groupe de risque faible bénéficiera d’un traitement peu agressif et moins toxique, alors qu’un patient avec une tumeur présentant des marqueurs de pronostic défavorable recevra d’emblée un traitement intensif. Deux approches sont actuellement privilégiées: 1) La recherche de la protéine transmembranaire (P-gp) responsable de la chimio-résistance multiple (MDR) chez les enfants atteints de LLA, et 2) la détection de la maladie résiduelle minime chez les patients atteints de neuroblastome.

La chimio-résistance, intrinsèque ou acquise, en particulier celle causée par la surexpression des protéines de transport transmembranaire, est l’une des causes majeures de l’échec thérapeutique chez les enfants souffrant de LLA. Plusieurs mécanismes de reflux conduisant à une résistance multiple aux drogues ont été décrits (entre autres P-gp, MRP1, BCRP et LRP), mais leur signification clinique chez les patients avec LLA est encore controversée. L’absence de standardisation des techniques de détection ne permet pas la comparaison des différentes études. De plus, la sensibilité des méthodes actuelles est encore insuffisante pour détecter de rares cellules tumorales à faible niveau de résistance. C’est pourquoi un objectif majeur de ce travail a été le développement, l’évaluation et l’optimisation d’une méthode sensible et spécifique de mesure de l’activité des P–gp. Sa relevance clinique chez les enfants souffrant de LLA a été évaluée dans une étude prospective et une synthèse de l’état des connaissances actuelles sur la valeur des autres protéines de chimio-résistance multiple a été réalisée. Plusieurs études ont montré un lien entre l’expression et/ou l’activité de P-gp, LRP, BCRP et MRP3 et une évolution défavorable de la maladie chez les enfants souffrant de LLA. Cependant des investigations plus poussées doivent encore confirmer ces résultats.
La détection fiable d’une infiltration médullaire par des cellules de neuroblastome représente au diagnostic une donnée essentielle pour établir le bilan d’extension et évaluer le pronostic, et ensuite pour mesurer la réponse au traitement. De plus le screening des préparations d’aphérèse est indispensable afin d’assurer une réinfusion de cellules souches exemptes de cellules tumorales, qui dans le cas contraire pourrait conduire à une rechute. L’analyse cytomorphologique classique de la moelle osseuse n’est informative que dans les cas d’invasion massive de la moelle au diagnostic ou lors de rechute. En revanche, le manque de sensibilité de cette technique ne permet pas la détection de micrométastases ou de la maladie résiduelle minime. Il est donc urgent de disposer de méthodes sensibles et spécifiques. Nous avons développé et standardisé des analyses sensibles faisant appel à des techniques d’immunocytochimie, de cytométrie de flux, de PCR quantitative en real-time. En plus, les résultats sur la valeur clinique des micrométastases chez des enfants souffrant de neuroblastome ont été analysés. En effet, plusieurs chercheurs ont démontré qu’une maladie résiduelle minime pouvait être associée à une probabilité élevée de rechute. Des études internationales et standardisées seront nécessaires pour confirmer ces résultats préliminaires.

En conclusion, les méthodes de détection développées dans le cadre de cette thèse ont permis la mise en évidence précise et sensible des phénomènes de chimio-résistance multiple et de maladie résiduelle minime chez les enfants atteints de cancer. Il pourrait en résulter une meilleure évaluation du risque évolutif et une anticipation thérapeutique avec une amélioration de la survie à long terme chez les enfants atteints de cancer.
Abbreviations

13-cis-RA: 13-cis retinoic acid
ABC: ATP-binding cassette
AHSCT: autologous hematopoietic stem cell transplantation
ALL: acute lymphoblastic leukemia
AML: acute myeloid leukemia
B-ALL: precursor B-cell lymphoblastic leukemia
BCRP: breast cancer resistance protein
CD: cluster of differentiation
CNS: central nervous system
CPC: criteria positive cell
CT-scan: computed tomography scan
c: cytoplasmic
del: deletion
dmin: double minute chromosomes
DNA: deoxyribonucleic acid
EFS: event-free survival
FAB: French-American-British
GALGT: GD2 synthetase (β1,4-N-acetylgalactosaminytransferase)
HSR: homogeneously staining regions
HVA: homovanillic acid
INPC: International Neuroblastoma Pathology Committee
INSS: International Neuroblastoma Staging System
JC-1: 5,5',6,6'-tetrachloro-1,1',3,3'-tetraethylbenzimidazolcarbocyanine iodide
LDH: lactate dehydrogenase
LOH: loss of heterozygosity
LRP: lung resistance protein
Mb: megabases
MDR: multidrug resistance
MIBG-scan: metaiodobenzylguanidine scan
MKI: mitosis-karyorrhexis index
MRD: minimal residual disease
MRI-scan: magnetic resonance imaging scan
mRNA: messenger RNA
MRP: multidrug resistance-associated protein
Abbreviations

NAG: neuroblastoma amplified gene
NC: negative cell
NCIC: not convincingly interpretable cell
PAS: periodic acid Schiff
PBSC: peripheral blood stem cell
P-gp: P-glycoprotein
QPCR: real-time quantitative RT-PCR
rho123: rhodamine 123
RNA: ribonucleic acid
RT-PCR: reverse transcriptase-polymerase chain reaction
SIOPEN: International Society of Pediatric Oncology European Neuroblastoma group
sP-gp: sister of P-glycoprotein
T-ALL: precursor T-cell lymphoblastic leukemia
TAP: transporter associated with antigen processing
TCR: T cell receptor
TdT: terminal deoxynucleotidyl transferase
TH: tyrosine hydroxylase
VMA: vanillylmandelic acid
WHO: World Health Organization
Dankwoord

Een dankwoord schrijven is eigenlijk het leukste wat er is. Het betekent immers dat mijn doctoraatsproefschrift bijna af is en ik eindelijk de mensen die mij de afgelopen jaren gesteund hebben in het zonnetje kan zetten.

In de eerste plaats wil ik mijn promotoren Prof. Dr. Jan Philippé en Prof. Dr. Geneviève Laureys bedanken. Zij hebben, samen met Prof. Dr. Yves Benoit, Prof. Dr. Catharina Dhooge en Dr. Barbara De Moerloose, het project in goede banen geleid en mij jarenlang uitstekend begeleid. Jan, jouw vertrouwen, enthousiasme, geduld en hulp hebben mij gevormd tot de onderzoeker die ik nu ben. Ik wil je bedanken voor de vele leerrijke discussies en je goede raad. Geneviève, jij hebt me enkele jaren geleden in de boeiende neuroblastoom-wereld binnengeleid. Dankzij jou kreeg ik de kans om met veel internationale wetenschappers samen te werken. Hartelijk dank! Barbara, dank je wel voor je steun en vriendschap. Je had altijd tijd voor mij en wist me steeds weer te motiveren. Ik hoop dat onze samenwerking hier niet eindigt,…

I like to thank the members of the jury, Prof. Dr. Geert Leroux-Roels, Prof. Dr. Andy Pearson, Prof. Dr. Jean Bénard, Prof. Dr. Marc Bracke, Prof. Dr. Frank Speleman, Prof. Dr. Lucien Noens and Prof. Dr. Yves Benoit for the critical evaluation of my thesis and the constructive comments.

Gedurende mijn 5 jaar op de dienst Klinische Biologie heb ik ontzettend veel geleerd. Ik zou dan ook alle stafleden willen bedanken voor de goede samenwerking. Ik heb er ook veel assistenten leren kennen. Ik wil hen danken voor hun hulp, interesse en bemoedigende woorden. Tussen het harde werken door, heb ik met hen veel leuk en ontspannende momenten beleefd. Even belangrijk waren de laboranten van het labo Hematologie, het Centrum voor Moleculaire Diagnostiek en de ‘weefselkweekers’ van de dienst Medische Genetica. In het bijzonder wil ik Detef Claeys, Lidia Vanballenberghae, Mieke Platteeuw, Elke Verbraekel en Jeannine Geerinck bedanken. Detlef, dank je wel voor alle praktische hulp. Zonder jou was ik verdronken in een onoverzichtelijke stapel cytopsins en MDR stalen.

Ik wil ook alle pediatrische oncologen bedanken voor de collectie van de stalen en de interesse in mijn werk. Dank ook aan alle patiënten en hun ouders voor de deelname aan de studies.
Dankwoord

Also many thanks to the members of the SIOP European Neuroblastoma Bone Marrow and RT-PCR subcommittees. This multicenter collaboration resulted in a publication included in this thesis. I enjoyed working with you and hope we will be able to continue our cooperation in the future.

Special thanks to Dr. Klaus Beiske for the instructive discussions, ideas, and help with the pictures. Your enthusiasm and interest in my work have been an important stimulus. I hope we will continue to meet and exchange ideas in the future,...

Dr. Nicole Gross and Apr. Katia Vandenbulcke, ‘merci beaucoup’ for your help with the French ‘résumé’.

Dank aan alle mensen die op welke wijze dan ook betrokken zijn bij de financiering van het in dit proefschrift beschreven onderzoeksproject. De grootste bron van financiering was het Instituut voor de Aanmoediging van Innovatie door Wetenschap en Technologie in Vlaanderen (IWT), maar ook het Kinderkankerfonds onder leiding van Prof. Dr. Yves Benoit heeft een belangrijke bijdrage geleverd. Hartelijk dank!

Tenslotte wil ik mijn ouders en broer Wouter bedanken voor de steun, de interesse in mijn werk en de nodige hulp wanneer het even wat moeilijker ging. ‘Weten’-schap werd me reeds met de paplepel ingegeven. Als kleuter speelde ik al met papa’s oscilloscoop en mama beantwoordde geduldig de eindeloze rij vragen: ‘Maar mama, waaróóóm?’ Wouter, jij wist me met je eindeloze fantasie altijd op te vrolijken en droeg er zorg voor dat ik niet vergat dat er ook nog andere dingen belangrijk waren in het leven.

Allemaal heel erg bedankt!

Katrien
Curriculum Vitae

ir. Katrien Swerts
Department of Pediatric Hematology and Oncology
Ghent University Hospital
De Pintelaan 185
9000 Ghent
Belgium
Tel: ++32/9/240.66.42
Fax: ++32/9/240.49.85
E-mail: Katrien.Swerts@UGent.be

EDUCATION

1989-1995 Sciences-Mathematics (8h), Regina Pacis Institute, Hove, Belgium
1995-1997 Bio-engineer (first cycle),
    Faculty of Science,
    University of Antwerp, Antwerp, Belgium
1997-2000 Bio-engineer in Cell and Gene Biotechnology,
    Faculty of Agricultural and Applied Biological Sciences,
    Ghent University, Ghent, Belgium (magna cum laude)
    Thesis: 'Multidrug resistance: a challenging search for new detection
    techniques' promoter Prof. Dr. G. Leroux-Roels, co-promoter Prof. Dr. J.
    Philippé
2000-2005 PhD training in Medical Sciences, Ghent University, Ghent, Belgium

PROFESSIONAL RECORD

2000-2004 Ph.D. candidate with grant of the Institute for the Promotion of Innovation
    through Science and Technology in Flanders (IWT-Vlaanderen)
2001-  Member of the SIOP European Neuroblastoma Bone Marrow Subcommittee
    (SIOP: International Society of Pediatric Oncology)
2002-  Secretary of the SIOP European Neuroblastoma Bone Marrow Subcommittee
2002-  Member of the SIOP European Neuroblastoma RT-PCR Subcommittee

LIST OF PUBLICATIONS

Publications in International Journals with Referee System

The combined analysis of P-glycoprotein expression and activity predicts outcome in
childhood acute lymphoblastic leukemia
Pediatric Hematology and Oncology 2003, 20: 381-391.


Publications in National Journals with Referee System


LIST OF ORAL PRESENTATIONS


Swerts K, De Moerloose B, Dhooge C, Laureys G, Benoit Y and Philippé J.
Lack of prognostic value of in vitro measurement of apoptotic response to glucocorticoids in childhood acute leukemia

Swerts K, De Moerloose B, Dhooge C, Laureys G, Benoit Y and Philippé J.
Detection of minimal residual disease in neuroblastoma patients
Clinical Biology Staff Meeting, Ghent, Belgium, 14/02/2002.

Swerts K, De Moerloose B, Dhooge C, Laureys G, Benoit Y and Philippé J.
Lack of prognostic value of in vitro measurement of apoptotic response to glucocorticoids in childhood acute leukemia

Swerts K, De Moerloose B, Dhooge C, Benoit Y, Laureys G and Philippé J.
Detection of residual neuroblastoma cells in bone marrow: comparison of three different detection techniques

Swerts K, De Moerloose B, Dhooge C, Benoit Y, Philippé J and Laureys G.
Detection of neuroblastoma cells by means of real-time quantitative RT-PCR
35th Congress of the International Society of Pediatric Oncology (SIOP 2003), Cairo, Egypt, 08-11/10/2003.

Swerts K, De Moerloose B, Dhooge C, Benoit Y, Philippé J and Laureys G.
MRD analysis in neuroblastoma (invited speaker)
Focus Conference Hematology, Anderlecht, Belgium, 14/11/2003.

Swerts K, De Moerloose B, Dhooge C, Benoit Y, Philippé J and Laureys G.
Multiparameter flow cytometric detection of residual neuroblastoma cells

Swerts K, De Moerloose B, Dhooge C, Benoit Y, Philippé J and Laureys G.
Minimal Residual Disease in Neuroblastoma Patients: a challenging search for new detection techniques
Pediatric Oncology Staff Meeting, Ghent, Belgium, 20/04/2004.

Swerts K, De Moerloose B, Dhooge C, Benoit Y, Laureys G and Philippé J.
Four-color flow cytometric detection of residual neuroblastoma cells

Swerts K, De Moerloose B, Dhooge C, Benoit Y, Philippé J and Laureys G.
Detection of residual neuroblastoma cells using a new four-color flow cytometric assay
Advances in Neuroblastoma Research, Genova, Italy, 16-19/06/2004.

Swerts K, De Moerloose B, Dhooge C, Benoit Y, Laureys G and Philippé J.
Detection of residual neuroblastoma cells in bone marrow using four-color flow cytometry
Detection of residual neuroblastoma cells using immunocytology/genetics (invited speaker)

Swerts K and Philippé J.
Multiparameter DNA-analysis: PI versus DRAQ5 (invited speaker)

LIST OF POSTER PRESENTATIONS (Only first author)

Swerts K, De Moerloose B, Dhooge C, Laureys G, Benoit Y and Philippé J.
Evaluation of a functional assay using JC-1 to evaluate P-gp expression in acute leukemia
17th General Meeting of the Belgian Hematological Society, Brussels, 01-02/02/20002.

Swerts K, De Moerloose B, Dhooge C, Laureys G, Benoit Y and Philippé J.
Detection and quantitation of residual neuroblastoma cells in bone marrow by real-time reverse transcription-PCR for tyrosine hydroxylase mRNA

The quantification of MDR1 expression by real-time RT-PCR has no prognostic value in childhood acute lymphoblastic leukemia
18th General Meeting of the Belgian Hematological Society, Brussels, Belgium 31/01/2003-01/02/2003.

Lack of prognostic value of quantitative MDR1 analysis by real-time RT-PCR in childhood ALL

Swerts K, De Moerloose B, Dhooge C, Laureys G, Benoit Y and Philippé J.
Multiparameter flow cytometric detection of residual neuroblastoma cells

Swerts K, De Moerloose B, Dhooge C, Laureys G, Benoit Y and Philippé J.
Comparison of two functional flow cytometric assays to assess P-gp activity in acute leukemia

Swerts K, De Moerloose B, Dhooge C, Benoit Y, Philippé J and Laureys G.
Detection of neuroblastoma cells by means of real-time quantitative RT-PCR
Advances in Neuroblastoma Research, Genova, Italy, 16-19/06/2004.

Swerts K, De Moerloose B, Dhooge C, Benoit Y, Philippé J and Laureys G. Detection of residual neuroblastoma cells using quantitative real-time RT-PCR
Advances in Neuroblastoma Research, Genova, Italy, 16-19/06/2004.


Swerts K, De Moerloose B, Dhooge C, Benoit Y, Laureys G and Philippé J. Detection of residual neuroblastoma cells in bone marrow using quantitative real-time RT-PCR


AWARDS

Grant for Young Researchers

Co-laureate of the Award ‘Prof. Dr. C. Hooft’, Period 2000-2003 (With Dr. B De Moerloose)

COURSES

‘Vormingsdagen Belgisch EMBnet Knooppunt’

Training Course in Hematology, Belgian Hematology Society
Belgium, December 2001 - May 2003.