Prognostic markers in acute myeloid leukemia
- A candidate gene approach
Ingrid Jakobsen
Prognostic Markers in Acute Myeloid Leukemia
A Candidate Gene Approach

Ingrid Jakobsen

Division of Drug Research
Department of Medical and Health Sciences
Linköping University, Sweden
Linköping 2018
©Ingrid Jakobsen, 2018

Cover illustration modified from iStock.com/kostenkodesign

Published articles have been reprinted with the permission of the copyright holder.

Printed in Sweden by LiU-Tryck, Linköping, Sweden, 2018

ISSN 0345-0082
I have no idea where this will lead us, but I have a definite feeling it will be a place both wonderful and strange.

_Special Agent Dale Cooper_
_Twin Peaks_
CONTENTS

ABSTRACT ......................................................................................................................... 1
POPULÄRVETENSKAPLIG SAMMANFATTNING ............................................................ 3
LIST OF PAPERS ................................................................................................................. 5
ABBREVIATIONS ................................................................................................................. 7
INTRODUCTION .................................................................................................................. 9
  Acute myeloid leukemia ................................................................................................. 9
  The difficult risk assessment ......................................................................................... 12
  AML treatment ............................................................................................................... 14
    Standard chemotherapy .............................................................................................. 14
    Targeted treatments .................................................................................................. 15
Evaluation of treatment response ...................................................................................... 15
What is pharmacogenetics and why does it matter? ......................................................... 16
The role of ABCB1 in AML treatment ............................................................................ 17
CDA, dCK and cN-II in Ara-C metabolism ...................................................................... 20
The role of IDH1/2 enzymes in AML ............................................................................ 22
AIMS ................................................................................................................................. 25
METHODS ......................................................................................................................... 27
  Overview ....................................................................................................................... 27
  Patient cohorts .............................................................................................................. 28
    Paper I ....................................................................................................................... 28
    Paper II ..................................................................................................................... 28
    Paper III .................................................................................................................... 29
    Paper IV .................................................................................................................... 29
Genotyping ....................................................................................................................... 29
  ABCB1 and DCK SNP genotyping using Pyrosequencing ............................................. 29
  Taqman SNP genotyping for CDA and cN-II variants ................................................ 33
  IDH mutation and IDH1 codon 105 SNP analysis ......................................................... 33
  In vitro cytotoxicity assay ............................................................................................ 33
RESULTS AND DISCUSSION

Genotype frequencies

ABCB1 SNPs

CDA, DCK and cN-II SNPs

IDH1/2 mutation and IDH1 codon 105 SNP

ABCB1 SNPs as prognostic markers in AML

ABCB1 variants, in vitro drug sensitivity and patient survival

1236C>T and 2677G>T in FLT3 subgroups

Does inconclusive mean irrelevant?

SNPs in genes related to Ara-C metabolism: Utility as markers of prognosis

Relationships with overall survival

In vitro data

Prognostic value of IDH2 mutations and IDH1 codon 105 SNP

IDH1 and IDH2 mutations

IDH1 codon 105 SNP

Risk-adapted therapy based on IDH status

CONCLUSIONS

FUTURE ASPECTS

ACKNOWLEDGEMENTS

REFERENCES

APPENDIX
Abstract

The standard treatment of acute myeloid leukemia (AML) consists of induction chemotherapy, most commonly daunorubicin together with the nucleoside analogue cytarabine (Ara-C), followed by consolidation chemotherapy and in selected cases allogenic stem cell transplantation (allo-SCT). Despite a high initial response rate, a considerable proportion of all AML cases eventually suffer from relapse and the five-year overall survival rate in patients >60 years is only around 15%. Based on cytogenetic analysis, patients are divided into low risk, intermediate risk, and high-risk groups. While low risk patients have a high chance of reaching and remaining in remission after standard induction therapy, high-risk patients are likely to suffer from relapse and should be scheduled for allo-SCT when first complete remission is reached. The intermediate risk group consists of normal karyotype (NK) patients and those with karyotypes of uncertain clinical relevance, but the outcomes are heterogeneous. In NK-AML patients, risk classification has improved with the addition of molecular markers including FLT3 internal tandem duplications (ITD) and mutations of NPM1 and CEBPA. Despite this development, there is a group of patients lacking reliable prognostic markers and in some cases the outcomes predicted do not match the outcomes observed, highlighting the need for additional markers. ABCB1 encodes a transporter protein responsible for the extrusion of cytotoxic compounds, including daunorubicin, over the cell membrane, and is a known resistance mechanism. Ara-C is subject to both activating and inactivating metabolic enzymes including DCK (activating), CDA and cN-II (inactivating). ABCB1, DCK, CDA and cN-II are all polymorphic, and SNPs affecting enzyme function and/or activity have potential as prognostic markers. In addition, recurrent IDH1/2 mutations lead to the expression of an enzyme with neomorphic activity associated with epigenetic alterations and disturbed differentiation. Mutations as well as a SNP in codon 105 of IDH1 have prognostic implications in AML, although the effects of different IDH mutations have been unclear. The aim of this thesis was to investigate SNPs in ABCB1 and genes associated with Ara-C metabolism, mutations in IDH1/2 and the IDH1 SNP, and their associations with treatment response and survival in AML. We show that the 1236C>T and 2677G>T SNPs in ABCB1 influence in vitro sensitivity towards AML drugs, with corresponding effects on NK-AML patient survival. These survival differences were seen mainly in patients lacking FLT3-ITD, further adding to the risk stratification. In contrast, the CDA SNPs 79A>C and -451C>T
appear to influence survival mainly in FLT3-ITD positive cases. In conclusion, the above-mentioned SNPs have the potential to add important information to risk classifications especially in NK-AML patients with the ambiguous FLT3-ITD-/NPM1- or FLT3-ITD+/NPM1+ genotypes. In addition, we have shown that IDH2 R140 mutation is associated with impaired survival in AML, and that the IDH1 codon 105 SNP appears to confer a worse outcome in a subset of intermediate risk patients without FLT3-ITD. With the introduction of next generation sequencing into clinical diagnostics, IDH mutations may not only provide prognostic information but also guide the selection of patients for new drugs targeting the variant enzyme. Our results indicate that in addition to leukemia-specific mutations, constitutional SNPs may prove useful for further individualizing care-taking and should be considered when implementing these new techniques.
möjligt att analysera stora delar av arvsmassan samtidigt är på väg att införas vid de större sjukhuslaboratorierna, och fokus ligger på att identifiera mutationer hos cancercellerna. Våra resultat visar att det inte bara är viktigt att titta på cancercellernas olika mutationer, utan att även vanligt förekommande medfödda genetiska variationer kan ha betydelse för patienternas prognos. Därför bör man överväga att även inkludera sådana medfödda variationer i analyserna vid diagnos.
LIST OF PAPERS

The following papers are included in this thesis, referred to in the text by their roman numerals (I-IV).


Other co-authored papers published during the Ph.D-period:

**Haematology**


**Ovarian cancer**


**Forensic sciences**


ABBREVIATIONS

ABCB1  ATP binding cassette subfamily B member 1
Allo-SCT  Allogenic stem cell transplantation
AML  Acute myeloid leukemia
APL  Acute promyelocytic leukemia
Ara-C  Cytosine arabinoside
Ara-CMP  Cytosine arabinoside monophosphate
Ara-CTP  Cytosine arabinoside triphosphate
ATP  Adenosine triphosphate
CDA  Cytidine deaminase
CEBPA  CCAAT/enhancer binding protein alpha
cN-II  Cytosolic 5'-nucleotidase II
CR  Complete remission
DCK  Deoxycytidine kinase
dNTP  Deoxynucleotide triphosphate
EFS  Event-free survival
ELN  European LeukemiaNet
FAB  French-American-British
FLT3  FMS-like tyrosine kinase 3
IDH  Isocitrate dehydrogenase
ITD  Internal tandem duplication
MDS  Myelodysplastic syndrome
NGS  Next generation sequencing
NK  Normal karyotype
NPM1  Nucleophosmine 1
OS  Overall survival
PCR  Polymerase chain reaction
PFS  Progression-free survival
PPi  Pyrophosphate
SNP  Single nucleotide polymorphism
TET2  Tet methylcytosine deoxyxygenase 2
α-KG  α-ketoglutarate
2-HG  2-hydroxyglutarate
INTRODUCTION

Acute myeloid leukemia

Acute myeloid leukemia is a blood cell malignancy characterized by acquired or, in rare cases, inherited genetic alterations leading to failed differentiation and increased proliferation of haematopoietic myeloid progenitors. Accumulation of myeloid blasts in the bone marrow will lead to suppression of normal haematopoiesis and to an abnormal peripheral blood picture. In around 25% of patients, the leukemia is preceded by a chronic bone marrow disorder like myelodysplastic syndrome (MDS), or by previous treatment with radiation or chemotherapy, but in most of the cases the aetiology is largely unknown. Around 300-350 patients are diagnosed with AML in Sweden each year, and although the median age is high (>70 years) all age groups are affected (1).

The disease often has a rapid course with a short time between the first symptoms and diagnosis, and if left untreated death usually occurs within a few weeks. The five-year overall survival (OS) is approximately 50-60% in patients <50 years, 20-40% in patients between 50 and 70 years, and less than 10% in those above the age of 70 (1, 2). Outcomes vary within age groups due to leukemia-specific variations but also due to patient-related factors such as comorbidity and performance status.
Figure 1. Bone marrow smear from an AML-Mo (acute minimally differentiated leukemia) patient. The blasts in this case are without clear granulation or the presence of needle-like structures called Auer rods. As indicated by the arrowheads, the blasts vary in size, have finely distributed chromatin and lack signs of obvious myeloid differentiation. The specific diagnosis needs to be confirmed with immune phenotype and molecular marker analysis, and the leukemia-specific immune phenotype profile can later be used for measurable residual disease monitoring. Arrow showing a small normal lymphocyte. Some mitotic events are visible, indicating proliferative activity (star). Image courtesy of Dr. Claes Malm, Linköping University Hospital.

Common symptoms preceding the diagnosis include general malaise, fatigue, nocturnal sweating, low-grade fevers, abnormal bruising or bleeding, or recurrent or persistent infections. Although increased peripheral white blood cell count with a high number of blasts is common, patients may also present with normal or low cell counts. In most cases deviations in differential counts are present, raising the suspicion of haematological malignancy. Bone marrow smears will also show deviating differential counts, and blast cells of varying morphology and different degrees of myeloid differentiation. In the case of an AML-Mo (acute minimally differentiated leukemia), shown in Figure 1, light microscopy of the bone marrow smear will indicate an acute leukemia, but further
analyses are required to definitively characterize it as myeloid. Diagnosis is confirmed by cytogenetic analysis of bone marrow aspirates or biopsy (or, if bone marrow sampling fails, peripheral blood) to identify chromosomal aberrations consistent with leukemia, fluorescence in situ hybridization (FISH) analysis for specific diagnostic/prognostic markers, flow cytometry analysis of cytoplasmic or cell surface markers, and molecular analysis of a selected set of genetic mutations of prognostic relevance. Cytogenetic and molecular analysis are important in the diagnostic work-up also for the classification of patients into low-, intermediate-, or high-risk groups, generally referring to the risk of treatment failure and/or relapse.

Previously, AML was classified based on morphology according to the French-American-British (FAB) system. While this still is a useful system in some cases, AML is now classified according to the more comprehensive World Health Organization (WHO) Classification of Tumours of Haematopoietic and Lymphoid Tissues, which was last updated in 2016 with revisions to the classification of myeloid neoplasms and acute leukemia (3). This system classifies AML cases into the following categories:

1. AML with recurrent genetic abnormalities;
2. AML with myelodysplasia-related changes;
3. Therapy-related AML;
4. AML not otherwise specified (NOS);
5. Myelosarcoma;
6. Myeloid proliferation related to Down’s syndrome;
7. Acute leukemias of ambiguous lineage (mixed phenotype acute leukemia; MPAL)

The AML NOS incorporates the previous FAB system for classification based on morphology and maturation when the AML cannot be otherwise specified. In the latest update, the new entity Myeloid neoplasms with germ line predisposition was also added, to include the rare subgroup of patients with familial myeloid neoplasms associated with germline mutations.
Acute promyelocytic leukemia (APL) belongs to the first category above and is caused by chromosomal rearrangements involving the retinoic acid receptor alpha (RARA) gene on chromosome 17. Most commonly this is a t(15;17)(q24;q21) translocation, fusing RARA with the promyelocytic leukemia gene (PML). The PML-RARA fusion make this special case of leukemia sensitive to all-trans retinoic acid (ATRA), which induces terminal differentiation of the immature leukemic cells. APL is treated with ATRA together with arsenic trioxide in accordance with a national APL programme, and due to the unique characteristics and treatment protocol of this entity compared to other forms of AML it is not further included in the work of this thesis. This also applies to paediatric patients, who are managed and treated according to a separate national programme.

The wider introduction of Next Generation Sequencing (NGS) technologies in recent years has resulted in deeper knowledge of the mutational landscape underlying the pathogenesis of AML (4, 5). Compared to many other cancers in adults, most AML cases have relatively few recurrent mutations. Those relevant to pathogenesis can largely be categorized into DNA methylation related genes (DNMT3A, TET2, IDH1/2), chromatin modifiers (e.g. EZH2 and ASXL1), transcription factor fusions (such as APL-specific PML-RARA fusion), tumour suppressors (e.g. TP53), signalling genes (e.g. FLT3, KIT and KRAS/NRAS), cohesin complex genes (including STAG2 and RAD21), spliceosome complex genes, myeloid transcription factor genes (e.g. CEBPA), and mutations of the Nucleophosmin (NPM1) gene.

New classification systems for AML subtypes based on the composition of the likely causative mutations and co-mutations have been suggested, but so far there is no global clinical consensus on which system to implement, and risk assessment systems always have to be adaptable and continuously evaluated in relation to new emerging treatments.

The difficult risk assessment

The aim of categorizing AML into risk groups is to identify patient groups with variable response (chance of achieving CR) and risk of relapse after standard treatment, and to provide guidance in decisions on therapy intensity and allocation to allo-SCT; so-called risk adapted therapy. Allo-SCT currently remains the only option for curing AML in patients who have a high risk of relapsing after chemotherapy only, and it is of vital importance to identify patients in need of this intervention. The decision on allocation to early SCT (in remission after first course of chemotherapy) is based on genetic risk group, other leukemia-related prognostic factors at diagnosis, patient-related risk factors such as high age, comorbidity and performance status, and response-related factors.
Generally, the risk profile is worse in older patients, with a higher frequency of high-risk cytogenetics, previous haematological disease like MDS, and comorbidities increasing the risk of therapy-related complications and early death, with a strong influence of performance status (6, 7). While those factors may favour palliative care, higher age does not necessarily disqualify a patient from intensive chemotherapy with curative intent (8). In addition, response related factors, i.e. how the patient responds to the first course of chemotherapy, influence the risk classification during the treatment of the patient.

Cytogenetic analysis (karyotyping) has traditionally been considered the most important tool for risk classification, and has been complemented during the last decade or so by molecular genetics analyses to further stratify patients into risk groups. Swedish national guidelines (9) for the diagnosis, risk classification and management of AML are based on experience and knowledge from clinical trials and data from the Swedish acute leukemia registry, and are in general coherent with the international recommendations from the European LeukemiaNet (ELN) published in 2010 (10).

Favourable risk patients are expected to have a good chance of reaching and staying in complete remission with standard chemotherapy only. Cytogenetic low-risk includes APL-related translocations involving RARA, and core binding factor (CBF)-AML with rearrangements between chromosomes 8 and 21 or within chromosome 16. In contrast, high-risk patients are predicted to have a very high risk of relapse after initial chemotherapy treatment. High-risk aberrations include deletions of chromosomes 5 or 7, chromosome 17p abnormalities (which are often associated with TP53 mutations), and complex karyotypes (harbouring three or more chromosomal aberrations). 40-50% of patients will be classified as cytogenetic intermediate risk, and although this intermediate group is considered to have an increased risk of relapse the outcome varies considerably.

In normal karyotype (NK) patients, and to some extent in karyotypes that cannot be classified as cytogenetic low or high risk, the analysis of CEBPA mutation, NPM1 mutation and internal duplications of FLT3 (FLT3-ITD) is currently used in clinical practice as a tool for risk classification in accordance with the 2010 ELN recommendations (10). In summary, the presence of double mutated CEBPA is considered a low risk marker; similarly, NPM1 mutated patients are classified as low risk but only in the absence of FLT3-ITD. The presence of FLT3-ITD in the absence of NPM1 mutation confers a high-risk profile. Still, there remains a group of patients with inconclusive cytogenetic aberrations or normal karyotype without any of the above-mentioned mutations, or normal karyotype
positive for both *NPM1* mutation and *FLT3-ITD*, highlighting the need for additional prognostic markers.

During the time frame of the writing of this thesis, some Swedish university hospital laboratories have started to implement NGS into AML diagnostics, most commonly investigating a panel of 20–50 genes known to be recurrently mutated in AML. It is likely that an NGS-based approach will also replace the current risk classification system where a selected set of aberrations are assessed using different methods, but the details of such an implementation are not yet specified in the national AML guidelines. It is encouraged that material is biobanked from both tumour and germ line source material, for future use.

### AML treatment

#### Standard chemotherapy

The Swedish national AML guidelines have been in use in clinical practice for less than 15 years, but the standard treatment for AML has remained roughly the same for over 40 years (11). Combination chemotherapy with the anthracycline daunorubicin and the nucleoside analogue cytosine arabinoside (also called cytarabine or Ara-C) is the backbone of AML treatment, followed by consolidation chemotherapy and/or allogenic stem cell transplantation (allo-SCT). Anthracyclines act mainly by topoisomerase II inhibition leading to DNA strand breaks, and by intercalating and inhibiting DNA and RNA synthesis (12, 13). Other mechanisms have been suggested, such as histone eviction from open chromatin (14). Ara-C in its active phosphorylated form cytosine arabinoside triphosphate (Ara-CTP) acts an antimetabolite inhibiting normal DNA synthesis (15).

The drugs are given as intravenous infusions over 3+5 days or 2+5 days for standard induction therapy and consolidation therapy; monotherapy with Ara-C for 5 days is administered as a last course of consolidation. Commonly the dosage is 60 mg/m² as an 8 h infusion for daunorubicin and 1 g/m² twice a day as 2 h infusions for Ara-C, but dose reductions are necessary in some patients and the recommendations have varied slightly over the years. Other drugs commonly used in AML treatment, specifically after initial treatment failure, include the intercalating and topoisomerase inhibiting agents amsacrine, mitoxantrone and etoposide, the purine analogue fludarabine, and the anthracycline idarubicin. As an example, induction treatment with amsacrine, cytarabine and etoposide may be preferred in patients with pre-existing cardiac conditions because of the cardiotoxicity associated with anthracycline treatment; alternatively, a prolonged infusion time may reduce the risk of heart failure (9, 16).
Similarly, dose reductions and prolonged infusion times of cytarabine are necessary in patients with a marked reduction in renal function to avoid toxicity (17).

**Targeted treatments**

While chronic myeloid leukemia is an example of huge success with the introduction of tyrosine kinase inhibitors targeting the CML-specific BCR-ABL1, such progress in the introduction of targeted treatments in AML has not yet been achieved. Targeted treatments recently approved or in clinical trial for the treatment of AML include the FLT3-inhibitor midostaurin for combination chemotherapy treatment in FLT3-mutated AML (18), the immunoconjugate gemtuzumab ozogamicin targeting CD33-positive AML (19), and the IDH2 inhibitor AG-221 in phase III clinical trial as a single agent in elderly patients (>60 years) with refractory or relapsed AML carrying IDH2 mutations (NCT02577406).

**Evaluation of treatment response**

A first evaluation of response can be done at day 15 after start of induction therapy to identify patients with a poor initial response. In Swedish clinical guidelines this is indicated by the presence of ≥10% blasts in the bone marrow (9). Early reinduction can be given, without awaiting bone marrow recovery, to decrease the risk of relapse and improve survival chances in those patients deemed fit to tolerate this second course. Although correlations between early response, remission rate and possibly long-term survival have been reported, there is no clear consensus on a precise bone marrow blast cut-off or clinical relevance depending on other risk factors (20).

Evaluation of complete remission (CR) should be performed at day 25-28 after start of induction therapy and is defined by the following criteria (9, 21):

- <5% blasts (bone marrow evaluation of at least 200 nucleated cells), no blasts with Auer rods
- Absence of extramedullary leukemia, no peripheral blasts
- B-neutrophil counts >1x10^9/L
- Thrombocyte particle count > 100x10^9/L
- Independence of erythrocyte transfusions

There are no minimum duration requirements for any of the above-mentioned criteria.
There is a clonal heterogeneity in AML, meaning that not all tumour cells harbour the same set of genetic aberrations and that a more resistant clone can expand after initial treatment and cause relapse in patients that meet the criteria for CR. This presents an obstacle to successful curative treatment. Proposed mechanisms include that the major clone harbours mutations making the cells more resistant to treatment, and they may also develop additional mutations and expand upon relapse. In addition, minor subclones (that may not be evident with standard diagnostic testing) may harbour a primary resistance to conventional therapy or evolve and acquire more mutations that can be selected and expand after the eradication of a major, sensitive clone (22, 23). Monitoring of measurable residual disease is routinely performed using analysis of leukemia phenotype markers by flow cytometry, or by analysis of leukemia-specific molecular markers. It aims at detecting any early signs of expansion of such remaining leukemic cells that lead to relapse, but the development of such methods is not covered further by the studies of this thesis.

What is pharmacogenetics and why does it matter?

Pharmacogenetics is the science of how inherited genetic variation can explain variability in drug distribution, metabolism and response to treatment. While all AML patients are connected by one thing – their diagnosis – they are also all unique individuals. They may differ in age, body composition and have different comorbidities. There are also other differences that are not visible on the surface, that can have a significant impact on how they respond to their treatment. Some differences may belong to the genetic makeup of the tumour cells, while others are attributed to constitutional genetic variations or epigenetic changes.

Along with technical developments and dropping analysis costs, the field has inevitably been moving towards a genome wide approach for identifying markers of treatment response and/or adverse drug reactions in cancer. In practice, this means analysing the whole haystack in the search for a small number of needles. One advantage is that prior knowledge of specific gene functions or a pre-existing hypothesis is not necessary for identification of variants that correlate to e.g. prognosis or drug resistance. The bias of selection of candidate genes of interest, based on the knowledge and assumptions of individual researchers, is also to a large extent avoided.

While this is an approach with the potential of discovering complex genetic relationships, it has also been associated with a relatively high false discovery rate and problems with reproducibility and interpretation of found associations; the opportunity to generate large amounts of sequencing data could potentially also confer an increased risk of losing the
functional biological anchoring. Why do we want to sequence on a very large scale? Is it always better or more informative to catch all the fish rather than a specific target? Technical standards and the development of best practice bioinformatics pipelines are continuously evolving, improving the quality and management of highly complex genetic data in cancer research, and likely that the increasing knowledge of the genetic landscapes of different diseases will lead to more specific applications of NGS in everyday clinic work.

In the work of this thesis, we took a hypothesis-based candidate gene approach, with a basis in knowledge of genes implicated in the distribution and metabolism of AML cornerstone chemotherapeutics. Single nucleotide polymorphisms (SNPs) in such genes have the potential to affect the expression, function and/or activity of the corresponding proteins, and may thereby also have an influence on both short- and long-term outcomes of treatment. Thus, candidate genetic variants may provide potential markers of prognosis that can be useful for individualization of treatment plans. In addition, as the genetic landscape of AML is being elucidated in large-scale sequencing studies, new treatments targeting specific recurrent mutations are emerging. Genes with recurrent mutations, as well as constitutional SNPs in such genes, may provide prognostic value and constitute markers for selecting patients most likely to benefit from treatment regimens including new targeting drugs.

The role of ABCB1 in AML treatment

The development of multidrug resistance in cancer is a clinically challenging problem that is the cause of a large proportion of cancer-related deaths. One of the first studies reporting multidrug resistance demonstrated that hamster cell lines grown in increasing concentrations of actinomycin D or daunorubicin to select for resistance, not only became resistant to the selection agent but also developed cross-resistance to other structurally diverse compounds (24). A few years later, a transporter named P-glycoprotein was identified in the resistant hamster cell lines (25) and the gene multidrug resistance 1 (MDR1) was subsequently reported to encode the human homologue (26, 27). This gene is also termed adenosine triphosphate (ATP) binding cassette (ABC) subfamily B member 1 (ABCB1) and is located on chromosome 7q21.12.

Like other ABC transporters, ABCB1 contains two intracellular ATP binding domains and two transmembrane domains responsible for substrate binding and transport under the hydrolysis of ATP. A structural comparison of ABCB1 to two other transporters, ABCG2 and ABCC1 (multidrug resistance-associated protein 1, MRP1), is shown in Figure 2a, and a schematic visualization of the transporter process is shown in Figure
While the nucleotide binding domains have similar structures and functions for the whole transporter family, the transmembrane domains have a high degree of heterogeneity which determines the variety of substrate that each member is able to transport across the cell membrane against the concentration gradient. ABCB1 is one of the most studied members of the ABC family, and is expressed at the apical surface of the epithelial cells of many human tissues including kidney, adrenal gland, liver and intestines (28). It is expressed in the hematological compartment in both peripheral blood lymphocytes and bone marrow, with higher expression in more undifferentiated progenitor cells (29).

Figure 2. a) Structural comparison of ABC transporters ABCG2, MDR1 (ABCB1) and MRP1 (ABCC1). B) Schematic visualization of ABCB1 transport. Adapted from Robey et al. 2018 (30) (reprinted with permission).

While the expression of ABCB1 in normal tissues can serve as a protective mechanism against exogenous toxins, it also confers a mechanism of resistance to antineoplastic agents in tumour cells when upregulated. The substrate specificity is wide, including compounds of diverse structures including anticancer drugs such as daunorubicin, etoposide, mitoxantrone, vincristine and paclitaxel (31). High expression of ABCB1 has been associated with impaired treatment outcomes and prognosis in different
cancers including acute leukemia (32, 33). The hypothesis is that increased drug efflux by the transporter leads to impaired treatment response due to suboptimal intracellular drug concentrations in the tumour cells. Early on, this implicated ABCB1 as a potential drug target, but clinical trials of transporter inhibitors have been unsuccessful; some explanations including toxicity or pharmacokinetic interactions resulting in the need for dose reductions of the primary chemotherapeutic treatment (30).

In addition to up-regulation, constitutional genetic variation such as SNPs may influence expression, substrate specificity and transporter activity in both normal and tumour cells. This results in the hypothesis that ABCB1 SNPs have potential as prognostic markers in AML treated with regimens including ABCB1 substrates. Given the extreme drug doses given during AML treatment, even small differences in genes related to pharmacokinetics of the drugs may have an impact on the outcome, in terms of both the effect on tumour cells and the susceptibility to severe toxic effects in normal cells. The most studied SNPs include the exon 12 1236C>T (synonymous), exon 21 2677G>T/A (893Ala>Ser/Thr) and exon 26 3435C>T (synonymous). These SNPs have a high frequency, are in linkage, and are present in the most common haplotype in Caucasian populations (34, 35).

The synonymous 3435C>T variant was the first SNP that was shown to have functional consequences on ABCB1 function in vivo, with reduced intestinal expression levels and increased plasma concentration of the substrate digoxin in healthy subjects homozygous for the alternative allele (36), and the T-allele was later shown to be associated with reduced mRNA stability (37). These first in vivo consequences demonstrated for the 3435C>T variant were later questioned by Kim et al. (2001), who showed a decreased plasma concentration of fexofenadine, another ABCB1 substrate, in subjects homozygous for the 3435T and 2677T alleles (38). Due to the genetic linkage it may be difficult to distinguish any individual effects of the three SNPs, and the effect appears at least in part to be substrate-dependent (39). Generally, the ABCB1 SNP research area has provided conflicting results regarding impact on expression, function and/or clinical consequences (40-48). A recent meta-analysis showed significant influence of 1236C>T, 2677G>T/A and 3435C>T on overall survival in AML patients with standard chemotherapy (49), an analysis partly based on the results of the work in this thesis.

Another non-synonymous SNP implicated to have functional and/or clinical relevance is the exon 11 1199G>A (400Ser>Asn) variant. Here too, the literature has been inconclusive, and substrate-specific influence of the variant have been reported for this SNP as well (45, 50-54).
CDA, dCK and cN-II in Ara-C metabolism

The nucleoside analogue Ara-C is transported into the cell primarily by the human equilibrative nucleoside transporter 1 (hENT1) (55, 56), where it then requires phosphorylation into the active metabolite Ara-C-triphosphate (Ara-CTP) (57). The rate limiting step is the conversion from Ara-C to Ara-C monophosphate (Ara-CMP). The gene for the responsible enzyme, deoxycytidine kinase (DCK), is located on chromosome 4q13.3 and altered expression or function of this enzyme may lead to poor outcomes and resistance to treatment (58, 59).

In addition to DCK, there are also enzymes responsible for deactivation of the drug. This includes cytosolic 5’-nucleotidase (cN-II), encoded by cN-II (also named NT5C2) on chromosome 10q24.32-q24.33, and cytidine deaminase (CDA), encoded on chromosome 1p36.12. cN-II dephosphorylates Ara-CMP while CDA deaminates Ara-C to the non-toxic metabolite uracil arabinoside (Ara-U); both enzymes inhibit the further production of Ara-CTP and thereby also limit the cytotoxic effect (60, 61).

Figure 3. Overview of Ara-C metabolic pathway, with key metabolic enzymes in yellow. The active metabolite Ara-CTP incorporates into DNA and RNA, blocking normal synthesis.
The relative balance in the activity of these enzymes may determine intracellular concentrations of the active metabolite Ara-CTP, which has been shown to vary significantly in AML patients, and thereby so too has the subsequent outcome of treatment (62-64).

Since the intracellular levels of Ara-CTP have been correlated to clinical response to Ara-C treatment (65), genetic variation in genes of importance for Ara-C uptake and metabolism may contribute to intraindividual differences in response and resistance. A few SNPs in the promoter of hENT1 have been reported to influence hENT1 expression, but the functional significance of the variants has not been clear, and the expressional differences reported have been attributed mainly to SNPs that are absent or very rare in Caucasian populations (66). Although deoxycytidylate deaminase (DCTD) has a role in Ara-CMP metabolism into the inactive Ara-UMP, no clear correlation between increased DCTD activity and Ara-C resistance has been shown (67); a SNP causing a significantly decreased enzyme activity have been identified but at a very low frequency in Caucasians (68).

The expression of DCK varies widely in leukemic cells and a decreased expression at relapse compared to diagnosis samples has been indicated (69). Several SNPs have been identified in the gene, and some have been associated with both DCK expression and Ara-CTP levels (70). This includes variants in the 3'-UTR, which are also representative of the most common haplotypes. The clinical relevance is not clear, but this indicates that SNPs in DCK may have potential as markers of response and/or prognosis in AML. Counteracting the production of Ara-CTP, the cytosolic 5'-nucleotidase 2 (cN-II) gene may also harbour SNPs with potential as clinical markers. Of the exonic non-synonymous SNPs identified, only the 7A>G (Thr3Ala, rs10883841) is common in Europeans with an allele frequency of approximately 13% in HapMap-CEU cohorts. Some associations between genetic variants of cN-II and expression and in vitro sensitivity towards Ara-C have been reported, but the clinical relevance of cN-II SNPs is largely unknown (71, 72).

Among CDA SNPs, the 79A>C (Lys27Gln, rs2072671) and the promotor variant -451C>T (rs532545) have a high frequency in Caucasians and have been associated with altered enzyme activity in vitro, treatment-related toxicity and prognosis in leukemia (68, 73-75). More recent research also supports a role for CDA polymorphisms in AML therapy outcome (76), but this was not clear at the initiation of the work in this thesis.
The role of IDH1/2 enzymes in AML

In the healthy cell, the IDH1 and IDH2 enzymes catalyse the decarboxylation of isocitrate into the unstable intermediate oxalosuccinate, while NADP+ is reduced to NADPH. In the following reaction carbon dioxide and α-ketoglutarate (α-KG) are produced. α-KG is a key intermediate in the citric acid cycle but has other functions as well, including nitrogen transport, and the tet methylcytosine deoxygenase 2 (TET2) enzyme requires this metabolite for 5-methylcytosine demethylation. The NADPH generation is also of importance for several functions including macromolecular synthesis and cellular defence systems against oxidative damage (77).

IDH1 is present predominantly in the cytosol and IDH2 in the mitochondrial matrix, and both enzymes have been shown to be mutated in cancer, with the earliest discoveries of IDH mutations in glioblastomas (78, 79). Soon after, both IDH1 and IDH2 mutations were also identified in leukemia, constituting one of the most common groups of mutations in normal karyotype AML (80, 81). In IDH1 codon 132 is affected and in IDH2 codon 140 and 172 are mutated; all three mutations are heterozygous and affect arginines in active sites of the enzymes (82). Not only do the mutations result in reduced oxidative decarboxylation of isocitrate, but the enzyme also acquires neomorphic activity. As a result, α-KG can be converted to the D-enantiomer of 2-hydroxyglutarate (2-HG), a metabolite that is normally present only at low levels but found to be elevated in many cases of NK-AML (81, 82).

![Diagram](image_url)

Figure 4. Wild type IDH1/2 catalyses the conversion of isocitrate to α-KG and CO₂ while NADP is reduced to NADPH. Mutant IDH1/2 acquires neomorphic activity converting α-KG to 2-HG in a reaction where NADPH is consumed. Image adapted from Cairns and Mak 2013 (77), reprinted with permission.
Although the normal function of 2-HG is largely unknown, mutant IDH is considered an oncogene with the 2-HG as an oncometabolite, influencing cell metabolism, proliferation and differentiation (83). For leukemia, this has been demonstrated by a correlation between high 2-HG levels and global hypermethylation phenotype with impaired differentiation (84). This was explained by the inhibition of the α-KG-dependent TET2 by 2-HG. In a similar fashion, alterations in histone methylation can occur through interference with α-KG-dependent histone demethylases (85).

The IDH genes have potential as clinical biomarkers in several ways. Firstly, IDH mutations may in themselves have an impact on prognosis, making genetic analysis a potential tool for identifying patients with increased risk of poor outcome after standard treatment and thus favouring early allocation to allo-SCT. Some studies on the impact of IDH mutations on prognosis in AML have been performed, with different conclusions. While Marcucci et al. (86) found an adverse impact of IDH1 R132 mutation in a subset of NK-AML patients, others could not confirm any impact on prognosis (87). This could be explained partly by differences in patient selection and subset analyses.

Secondly, in addition to the potential as prognostic markers, IDH mutations are likely to become markers of selection to new targeted drugs that are currently under development. From a health economics perspective, the introduction of new treatments to the market represents challenges in terms of resources; the drugs are often expensive and there is a need to prioritize the patients who are likely to benefit the most from such treatments, and to avoid overtreatment of patients who are unlikely to benefit. This could be investigated in drug trials that consider both IDH mutational status and perhaps also the presence of SNPs with implications for enzyme function and treatment outcome.

A synonymous SNP (rs11554137) affecting codon 105 with a GGC>GGT change, is located in the same exon as the IDH1 R132 mutation and present in the general population at a minor allele (T) frequency of around 5%. The functional consequence of this SNP has not been clearly determined, but hypotheses include altered stability of mRNA or effects on splicing, and differences in IDH1 expression have been demonstrated between genotypes (88, 89). It was also speculated that altered NADP+ production related to changes in IDH1 expression could have an influence on the haematopoiesis and thereby also chemosensitivity. Consequently, the IDH1 codon 105 SNP may constitute a potential marker of treatment outcome and prognosis in AML. A few studies have implicated a prognostic value of this SNP in leukemia patients, but the impact seems to be dependent on age and other factors such as cytogenetics and molecular markers (e.g. FLT3-ITD and NPM1 mutation) (88, 90, 91).
The overall aim of this thesis was to investigate the utility of candidate gene variants as markers of induction treatment response and prognosis in acute myeloid leukemia, with the long-term goal of better risk stratification and individualized treatment.

The specific aims of this thesis were to investigate:

- The relationship between single nucleotide polymorphisms of the drug transporter gene *ABCB1*, treatment outcome and resistance in NK-AML;
- *DCK*, *cN-II* and *CDA* SNPs and their relevance as markers of treatment response and long-term outcome in NK-AML;
- *IDH1/2* mutations and *IDH1* codon 105 SNP and their implications on treatment response and prognosis in unselected AML patients.
Overview

A brief overview of the patient cohorts in the papers of this thesis is presented below (Table I). For all studies, blood or bone marrow samples collected at diagnosis before the initiation of treatment were used for the analyses. Clinical data, including patient characteristics, cytogenetic evaluation, treatment response and survival, were retrieved from medical records to investigate genotype correlations to response (CR or no CR) and survival (OS and progression-free survival (PFS) or event-free survival (EFS)) in univariable- and multivariable statistical analyses. In patients where FLT3 and NPM1 mutational statuses were missing, additional analyses were performed where possible.

To compare genotype frequencies, two reference populations were used. These were a reference population of 400 individuals of comparable age and gender distribution, from the south-eastern region of Sweden (for Paper I and Paper III), and the HapMap-CEU population (for Papers II and IV).

All studies were performed with approval from regional ethics committees and in accordance with the Declaration of Helsinki.
Table I. Study overview.

<table>
<thead>
<tr>
<th>Paper</th>
<th>Number of patients</th>
<th>Cytogenetics</th>
<th>Time span</th>
<th>Participating centres</th>
<th>Candidate gene(s)</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>100</td>
<td>NK</td>
<td>1988-2007</td>
<td>Linköping Huddinge</td>
<td>ABCB1</td>
<td>Pyrosequencing, cytotoxicity assay</td>
</tr>
<tr>
<td>II</td>
<td>205</td>
<td>NK</td>
<td>1988-2009</td>
<td>Linköping Huddinge Uppsala Gothenburg</td>
<td>CDA dCK cN-II</td>
<td>Pyrosequencing, TaqMan SNP genotyping, cytotoxicity assay, methylation assay</td>
</tr>
<tr>
<td>III</td>
<td>201</td>
<td>NK</td>
<td>1988-2009</td>
<td>Linköping Huddinge Uppsala Gothenburg</td>
<td>ABCB1</td>
<td>Pyrosequencing</td>
</tr>
<tr>
<td>IV</td>
<td>189</td>
<td>Unselected</td>
<td>1988-2010</td>
<td>Linköping Huddinge</td>
<td>IDH1/2</td>
<td>Direct sequencing</td>
</tr>
</tbody>
</table>

NK: Normal karyotype

Patient cohorts

Paper I
Paper I (92) included samples from 100 adult patients (mean age 63 years) diagnosed with de novo NK-AML between 1988 and 2007 at Linköping University Hospital and Karolinska University Hospital in Huddinge, Sweden. No cases of secondary leukemias were included. In most patients, the treatment included anthracyclines or mitoxantrone in combination with Ara-C. Leukemic cells from 56 of the patients were available for in vitro cytotoxicity testing.

Paper II
In Paper II (93), 205 de novo NK-AML patients, mean age 59 years, from four different Swedish centres – Linköping, Huddinge, Gothenburg and Uppsala – were included. Patients were diagnosed between 1988 and 2009. All patients except two received induction treatment including Ara-C and treatment was given with curative intent. Until 2004, Ara-C was generally recommended at doses of 200 mg/m² as 24h infusions for 7 days, while
later national guidelines prescribed 1000mg/m\(^2\) in 2h intravenous infusions two times a day for 5 days. *In vitro* cytotoxicity results from Paper I were also included in Paper II.

**Paper III**
The patient cohort in Paper III (94) was based on the 100 NK-AML patients from Paper I, and further increased by 110 additional patients from Gothenburg and Uppsala, to enable a more detailed subgroup analysis. After data collection, nine patients not treated with curative intent were removed, resulting in a final cohort of 201 *de novo* NK-AML patients diagnosed between 1988 and 2009 with a mean age of 59 years. Patients diagnosed in 2005 or later were treated according to national guidelines (daunorubicin and Ara-C for 3+5 days), while patients diagnosed earlier were treated according to regional guidelines. Most commonly, these included Ara-C together with either daunorubicin or idarubicin for 7+3 days, with lower Ara-C doses compared to the current standard.

**Paper IV**
189 cytogenetically unselected AML patients (median age 64 years) from Linköping and Huddinge were analysed in Paper IV (95); 57% NK-AML, 40% with aberrant karyotype, and 3% with undetermined karyotype. 46% were categorized as intermediate risk; 17% were low risk and 29% were high risk. Risk assessment was missing in 8%. Treatment protocols varied depending on year of diagnosis (before or after 2005), but the most common regimens included anthracyclines in combination with Ara-C, with or without the additional drugs.

**Genotyping**

*ABCB1* and *DCK* SNP genotyping using Pyrosequencing
The *ABCB1* SNPs  1199G>A (Ser400Asn, rs2229109), 1236C>T (silent, rs1128503), 2677G>T/A (Ala893Ser/Thr, rs2032582) and 3435C>T (silent, rs1045642), and *DCK* SNP 3’-UTR 948T>C (rs4643786) were analysed by Pyrosequencing (Paper I-III). This is a method suitable for rapid SNP detection and sequencing of short DNA-sequences. The limit is around 200 nucleotides, and 50-60 nucleotides can routinely be sequenced in most templates with limitations related to PCR product quality, secondary structures in the template, and decreased enzyme efficiency with longer run times (96).
This is a sequencing-by-synthesis method based on the generation of pyrophosphate through the incorporation of nucleotides by a DNA polymerase into a single-stranded DNA template (97). In principle, before the pyrosequencing, a DNA region covering the SNP(s) of interest is amplified in a PCR reaction in which one of the customer-designed primers is biotinylated. Binding to streptavidin-coated sepharose beads enables capturing on filter probes, and with a washing process using sodium hydroxide as a denaturing agent, single-stranded DNA templates are prepared and transferred to a sequencing plate.

In this plate, a sequencing primer complementary to the region adjacent to the SNP is bound to the templates. Enzyme and substrate mixtures are added and, based on knowledge of the sequence, deoxynucleotide triphosphates (dNTPs) are then added in a predefined dispensation order. When a dNTP complementary to the strand is incorporated by the DNA polymerase, pyrophosphate (PPi) is released. The enzyme ATP-sulphurylase will convert the PPi into ATP in the presence of the substrate adenosine 5’-phosphosulphate, and the generated ATP will be used by luciferase to catalyse the conversion of luciferin to oxyluciferin. This light-generating reaction is detected by a camera, and the estimated amount of light is translated into a peak in a pyrogram. Before the addition of a new nucleotide, unincorporated dNTP and excessive ATP from the previous cycle will be degraded by apyrase.

The amount of light generated is proportional to the amount of incorporated dNTP, and thus the heights of the peaks can be used to determine the sequence and genotype of a variable position based on theoretical pyrogram outcomes. dNTPs known not to be present in a certain position are used as an internal negative control, and non-polymorphic regions serve as positive reference peaks. In addition, a standardized control oligo can be used as a control for sample preparation and run quality. An overview of the Pyrosequencing principles is shown in Figure 5, and example pyrograms for one of the SNPs analysed (ABCB1 3435C>T) are shown in Figure 6.
Figure 5. The principle of the pyrosequencing technique. A) Biotinylation of one of the PCR primers is utilized to prepare a single-stranded DNA template to which a sequencing primer is annealed. During the sequencing, each incorporation of a dNTP results in the generation of pyrophosphate (PPi). B) PPi is used by the enzyme sulfurylase to convert the substrate APS into ATP, which is further used by luciferase to convert luciferin into oxyluciferin, a light generation reaction that can be detected by a camera and visualized as a peak in a pyrogram. C) Between each nucleotide dispensation, excess dNTP and ATP from the previous cycle are degraded by apyrase. D) The height of each peak in the resulting pyrogram is proportional to the number of nucleotides incorporated into the growing strand, and thereby the sequence of the sample can be determined. (Image adapted from https://www.qiagen.com/se/resources/technologies/pyrosequencing-resource-center/technology-overview/?akamai-feo=off, printed with permission.)
Figure 6. Pyrograms corresponding to the three different $ABCB1\ 3435C>T$ genotypes, with the variable region highlighted: C/C (top), T/C (middle) and T/T (bottom). The height of each individual peak, corresponding to the amount of light generated, is proportional to the number of nucleotides in the sequence. Sequence to analyse (forward strand): GAT[T/C]GTG. E = enzyme mix dispensation, S = substrate mix dispensation. Two internal negative control dispensations are included at nucleotide dispensations no 1 (C) and no. 6 (A).
**Methods**

**Taqman SNP genotyping for CDA and cN-II variants**
The two CDA SNPs, 79A>C (Lys27Gln, rs2072671) and -451C>T (promotor variant, rs532545), and the cN-II SNP 7A>G (Thr3Ala, rs10883841) were evaluated using TaqMan SNP Genotyping Assays according to the manufacturer’s instructions but in a reaction volume of 10µL (Paper II). The assay kit contains a pair of unlabelled primers and two probes, one for each SNP variant, labelled in the 5’ end with fluorescent reporter dyes and in the 3’ end with a minor groove binder and non-fluorescent quencher. Genomic DNA, a genotyping master mix, primers and probes are added in a reaction mix, where the probes will anneal specifically to its complementary sequence in between the primers. During sequencing, the DNA polymerase will cleave probes that have hybridized to the target, thereby releasing the quencher leading to an increase in the fluorescence by the reporter dye. Depending on which reporter shows an increase in fluorescence, it can be determined which alleles are present. Compared to Pyrosequencing, where the analysis requires working with open PCR-products and thereby has an increased risk of contamination, the TaqMan analysis is performed in a closed system after the addition of the manufacturer-validated reagents. This decreases the contamination risk; however, one does not get the benefit of a pyrogram as a quality control.

**IDH mutation and IDH1 codon 105 SNP analysis**
Recurrent mutations of IDH1 and IDH2 were analysed in Paper IV, with detection of IDH1 codon 132, IDH2 codon 140 and 172 mutations, as well as the previously reported IDH1 codon 105 SNP (G105G, rs11554137). PCR reactions were performed for each gene and the product was purified using ExoSAP-IT, followed by direct sequencing. The resulting sequencing traces were compared to the reference IDH1/2 sequences (NM_005896.2 and NM_002168.2, respectively) to detect the genetic variants.

**In vitro cytotoxicity assay**
Leukemic cells from 56 NK-AML patients were available to investigate differences in *in vitro* sensitivity towards conventional chemotherapeutic drugs between genotypes in Papers I and II. The cells were incubated with drugs in concentrations chosen to mimic *in vivo* exposure (98, 99): Ara-C 0.5 µM (continuously), daunorubicin 0.2 µM, etoposide 20 µM and mitoxantrone 0.1 µM (during 1 h). After the short-time incubations, cells were spun down, supernatant removed, and fresh medium added. After 96 h incubation, a bioluminescence method was used as a marker of cell viability, measuring intracellular ATP concentration (100). The incubations were done in duplicate and with drug-free controls. Like the chemistry in Pyrosequencing, the bioluminescence method utilizes the
catalytic activity of the firefly luciferase, where the amount of light emitted during the enzymatic reaction is proportional to the ATP content.
RESULTS AND DISCUSSION

Genotype frequencies
All SNPs were found in expected frequencies compared to the reference populations, and in accordance with the Hardy Weinberg equation; the \textit{DCK} 3'-UTR *165C>T (rs4643786) homozygous C/C genotype is not present in European populations. Different frequencies of the \textit{IDH1/2} mutations in AML have been reported in previously published studies, depending on the composition of patient populations. Our data did not represent any extreme, with frequencies comparable with other mixed AML populations. Genotype frequencies for all variants (\textit{ABCB1}, \textit{CDA}, \textit{DCK}, cN-II, \textit{IDH1/2} mutations and \textit{IDH1} codon 105 SNP) are presented in the following sections together with frequencies for reference populations; for \textit{IDH} mutations, references are given to a selection of publications on AML populations.
ABCB1 SNPs

Due to the low frequency of the alternative 2677A allele, patients carrying this variant were excluded from the survival analyses. Genotype frequencies are presented in Table II.

Table II. ABCB1 SNP frequencies in AML and in a Swedish reference population. The frequencies did not differ significantly (p>0.05).

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype</th>
<th>AML patients* (n=202)</th>
<th>Swedish reference population (n=400)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1199G&gt;A</td>
<td>G/G</td>
<td>185 (91.6%)</td>
<td>362 (90.5%)</td>
<td>0.66</td>
</tr>
<tr>
<td>rs2229109</td>
<td>G/A</td>
<td>16 (7.9%)</td>
<td>37 (9.25%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A/A</td>
<td>1 (0.5%)</td>
<td>1 (0.25%)</td>
<td></td>
</tr>
<tr>
<td>1236C&gt;T</td>
<td>C/C</td>
<td>66 (32.7%)</td>
<td>133 (33.25%)</td>
<td>0.91</td>
</tr>
<tr>
<td>rs1128503</td>
<td>C/T</td>
<td>98 (48.5%)</td>
<td>187 (46.75%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T/T</td>
<td>38 (18.8%)</td>
<td>80 (20%)</td>
<td></td>
</tr>
<tr>
<td>2677G&gt;T/A</td>
<td>G/G</td>
<td>63 (31.2%)</td>
<td>124 (31%)</td>
<td>0.97</td>
</tr>
<tr>
<td>rs2032582</td>
<td>G/T</td>
<td>93 (46.0%)</td>
<td>184 (46%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T/T</td>
<td>38 (18.8%)</td>
<td>75 (18.75%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G/A</td>
<td>4 (2.0%)</td>
<td>10 (2.5%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A/T</td>
<td>4 (2.0%)</td>
<td>6 (1.5%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A/A</td>
<td>0 (0%)</td>
<td>1 (0.25%)</td>
<td></td>
</tr>
<tr>
<td>3435C&gt;T</td>
<td>C/C</td>
<td>39 (19.3%)</td>
<td>87 (21.75%)</td>
<td>0.53</td>
</tr>
<tr>
<td>rs1045642</td>
<td>C/T</td>
<td>98 (48.5%)</td>
<td>175 (43.75%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T/T</td>
<td>65 (32.2%)</td>
<td>138 (34.5%)</td>
<td></td>
</tr>
</tbody>
</table>

*Including patients from Paper I and Paper III.
CDA, DCK and cN-II SNPs

Genotype frequencies are presented in Table III. For the cN-II 7A>G SNP, genotyping was unsuccessful in three patients, probably because of poor sample quality.

Table III. Genotype frequencies for SNPs in CDA, DCK and cN-II. The frequencies did not differ significantly between AML patients and the reference population.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype</th>
<th>AML patients (n=205)</th>
<th>Reference population (n=99)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDA 79A&gt;C rs2072671</td>
<td>A/A</td>
<td>84 (41%)</td>
<td>44 (44.4%)</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>A/C</td>
<td>95 (46.3%)</td>
<td>42 (42.4%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C/C</td>
<td>26 (12.7%)</td>
<td>13 (13.1%)</td>
<td></td>
</tr>
<tr>
<td>CDA -451C&gt;T rs532545</td>
<td>C/C</td>
<td>83 (40.5%)</td>
<td>51 (51.5%)</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>C/T</td>
<td>101 (49.3%)</td>
<td>38 (38.4%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T/T</td>
<td>21 (10.2%)</td>
<td>10 (10.9%)</td>
<td></td>
</tr>
<tr>
<td>DCK 3’-UTR 948C&gt;T rs4643786</td>
<td>C/C</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>C/T</td>
<td>16 (7.8%)</td>
<td>12 (12.1%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T/T</td>
<td>189 (92.2%)</td>
<td>87 (87.9%)</td>
<td></td>
</tr>
<tr>
<td>cN-II 7A&gt;G rs10883841</td>
<td>A/A</td>
<td>138 (68.3%)</td>
<td>73 (76%)</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>A/G</td>
<td>60 (29.7%)</td>
<td>25 (23%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G/G</td>
<td>4 (2%)</td>
<td>1 (1%)</td>
<td></td>
</tr>
</tbody>
</table>

1The 1000 Genomes CEU population. 2In Paper II the C/T and T/T genotype frequencies were reported as 92.2% C/T and 7.8% T/T in supplemental Table II; an erratum has been sent to the American Journal of Hematology but was not in print at the conception of this thesis. 3Genotyping failed in three patients; frequencies calculated based on n=202.
**IDH1/2 mutation and IDH1 codon 105 SNP**

All *IDH1* mutations affected codon 132, with a total of 7.9% of the patients being mutated. *IDH2* mutations occurred in two hotspots, codon 140 and 172, in 13.8% of the patients. R140Q (arginine to glutamate) was the most common substitution in our cohort, and the mutations were mutually exclusive. In summary, *IDH* mutations affected 21.7% of the patients in our unselected AML cohort. The additional *IDH1* codon 105 SNP was identified at a frequency similar to those previously published on adult and paediatric AML.

Table IV. Frequencies of *IDH1/2* mutations and *IDH1* SNP in our study and a selection of references for previously reported frequencies in AML.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Variant</th>
<th>Unselected AML (Paper IV, n=189)</th>
<th>Previously reported in AML</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>IDH1</em></td>
<td>R132C</td>
<td>7 (3.7%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R132H</td>
<td>6 (3.2%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R132G</td>
<td>1 (0.5%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R132L</td>
<td>1 (0.5%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total R132</td>
<td>15 (7.9%)</td>
<td>5.5-10.4%</td>
<td>(80, 81, 87, 101-105)</td>
</tr>
<tr>
<td><em>IDH2</em></td>
<td>R140Q</td>
<td>20 (10.6%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R140G</td>
<td>1 (0.5%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total R140</td>
<td>21 (11.1%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R172K</td>
<td>5 (2.6%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total R140 &amp; R172</td>
<td>26 (13.8%)</td>
<td>6.1-17.7%</td>
<td>(81, 87, 102, 104-107)</td>
</tr>
<tr>
<td><em>IDH1/IDH2</em></td>
<td>All mutations above</td>
<td>41 (21.7%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>IDH1 (SNP)</em></td>
<td>Codon 105C&gt;T (rs11554137)</td>
<td>20 (10.6%)</td>
<td>11-11.7%</td>
<td>(88, 91)</td>
</tr>
</tbody>
</table>
**ABCB1 SNPs as prognostic markers in AML**

In Paper I and Paper III, SNPs in the drug transporter gene *ABCB1* were investigated. This was based on the hypothesis that since ABCB1 extrudes drugs commonly used in the treatment of AML, including the anthracycline Daunorubicin, polymorphisms implicated to alter ABCB1 function or expression are potential biomarkers of prognosis. Earlier, no one had, to our knowledge, investigated the influence of *ABCB1* SNP in normal karyotype AML while taking the clinically relevant *FLT3*-ITD and *NPM1* mutations into account. We determined the *FLT3*-ITD and *NPM1* mutational status in part of our 100-patient cohort (those diagnosed before the introduction of *FLT3*-ITD and *NPM1* analysis in clinical routine) and included these data in our analysis. As expected, patients with mutated *NPM1* appeared more likely to achieve CR after induction therapy, and *NPM1* mutated/*FLT3*-ITD negative patients also had a longer OS compared to other patients.

**ABCB1 variants, *in vitro* drug sensitivity and patient survival**

Among the *ABCB1* SNPs, only 1236C>T (rs1128503) and 2677G>T/A (Ala893Ser/Thr, rs20325829) showed any correlation to outcome, with better survival in patients carrying the alternative T alleles. This was also reflected in a higher *in vitro* sensitivity towards the ABCB1 substrates mitoxantrone and daunorubicin in cells from *FLT3*-ITD negative leukemic blasts from patients with 1236T/T or 2677T/T genotype. However, it must be considered that the cells were a small representation of the patient cohort and that the cytotoxicity assay displayed large variations.

In addition, the limited sample size did not allow investigation of the 1199G>A SNP, since only three patients with cells available for *in vitro* testing were heterozygous G/A in this position. In the eight patients carrying the 1199A allele, survival appeared to be impaired compared to the G/G genotype, and only one patient with very early death (before routine response evaluation of the bone marrow) was homozygous in this position. This patient was not included in the follow-up study in Paper III. A small study on ovarian cancer patients treated with the ABCB1 substrate paclitaxel has shown similar results to ours (108), and the G/A genotype has also been associated with adverse outcome with an increased risk of relapse in childhood ALL (53). Even though this SNP should not be completely dismissed as a marker of prognosis based on our results, it is relevant to consider whether the potential difference in outcome for NK-AML patients would motivate routine screening of this relatively rare SNP. In summary, our results from Paper I pointed towards the two more frequent SNPs 1236C>T and 2677G>T as markers of prognosis in NK-AML.
1236C>T and 2677G>T in FLT3 subgroups

To further investigate the significance of these two SNPs for the prognosis in different subgroups of NK-AML, the patient cohort from Paper I was expanded in a multicentre collaboration (Paper III). In the multivariable survival analyses, this meant that the (at the time) recently updated AML guidelines with risk stratification based on the combined FLT3-ITD/NPM1 mutation status could be applied. Since the in vitro results demonstrated differences in drug sensitivity in FLT3-ITD negative patient cell lines, survival was also analysed in subgroups based on FLT3 status with age, gender, NPM1 mutation, treatment and ABCB1 genotype as covariables. Analyses confirmed that the alternative alleles of 1236C>T and 2677G>T conferred a better prognosis compared to the wild type C/C and G/G genotypes. As expected, FLT3-ITD+/NPM1- patients had a significantly inferior outcome compared to the lower risk FLT3-ITD-/NPM1+ patient group. Our results also suggested allo-SCT as a factor associated with improved outcome regardless of FLT3-ITD status, although the lowest hazard ratio was observed in the FLT3-ITD+ subgroup.

Subgroup analysis revealed that the 1236C>T and 2677G>T variants predominantly had an impact on survival in FLT3-ITD- patients (Figure 7). Earlier publications have indicated that SNPs have a relevant influence on ABCB1 gene expression in patients with leukemia-specific induction of ABCB1, and that FLT3-IDT+ cases are less likely to co-express ABCB1 (109). Also, other differences between cells that express ABCB1 and cells with FLT3-ITD have been demonstrated, including WBC at diagnosis, growth rate, and differential expression of the transcriptional regulatory factor FOXO1 (110, 111). Unfortunately, we did not have the opportunity to measure ABCB1 expression in our patients, but theoretically a low or absent ABCB1 expression in FLT3-ITD+ cases provides a possible explanation as to why the genotype effect is negligible in this subgroup.

Since Paper III was an extension of the Paper I cohort, survival was also investigated separately in the new patients. The overall survival for the different genotypes followed the same pattern as for the entire cohort, but without statistical significance. Some differences can be noted between the cohorts, such as a higher median age in the earlier cohort (median age 65 vs 59 years) and the fact that a larger proportion of the newer patients were diagnosed after the introduction of new national guidelines for the management of AML. This was to some extent adjusted for in the main analysis by including year of diagnosis as a co-variable, but one must still consider the possibility that the influence of ABCB1 genotype might not be as prominent as indicated in our first publication.
Figure 7. Better overall survival in FLT3-ITD- (i.e. wildtype) NK-AML patients carrying the alternative ABCB1 1236T or 2677T allele. Median OS (A) 20 vs. 49 months, and (B) 20 vs 35 months.
Does inconclusive mean irrelevant?

Overall, previous studies on the prognostic influence of ABCB1 SNPs in different AML populations have been inconclusive. While Illmer et al. (44) showed poor outcomes in patients with the wild-type 1236C/C, 2677G/G and 3435C/C, consistent with our results, Kim et al. (48) contrastingly reported improved chance of CR and longer event-free survival for the 2677G/G and 3435C/C genotypes. Yet others have failed to demonstrate any significant differences in chance of response or survival associated with ABCB1 genotype (112, 113). Possible explanations for the diverging results include differences in patient populations regarding age or variations in treatment including the P-gp inhibitor tested in the study cohort of van der Holt (112). Regarding different origin of patient cohorts, this may be relevant due to differences between populations in haplotype structures and presence of other polymorphisms that were not analysed here.

However, one must consider the influence of ABCB1 SNPs from more than the leukemia cell perspective. During treatment, healthy cells in basically all tissues are exposed to the drugs, and differences in function and/or expression of the transporter in those cells may contribute to the sensitivity of the patient to severe treatment-related toxicity and mortality. This is the case especially in the treatment of AML; severe bone marrow toxicity is a sign that the treatment is effective on the leukemia, but it also means that the patients are exposed to a near-death situation with small margins. In such a case, a SNP causing decreased transport of a substrate could have two theoretical consequences: Firstly, a higher intracellular drug concentration in the tumour cells would increase the efficacy and potentially improve the chances of eliminating the cancer. Secondly, in all proliferative tissues the impaired transport would give similar increases in intracellular drug concentrations with a subsequent higher strain on the healthy cells. Thereby, inhibition of transporter function may not necessarily lead to improved outcomes in all patients treated with ABCB1 substrates.

It can also be speculated that an abnormally high ABCB1 expression in tumour cells could mask any subtle effect of a SNP, or that mutations in the gene would be of more importance. Sequencing data from The Cancer Genome Atlas shows increased expression of ABCB1 in around 13% of AML tumour samples, and mutations or rearrangements affecting the transporter function/expression appear to be very rare or absent (as reviewed by Robey et al. 2018 (30)). In particular, gene rearrangements leading to increased expression through promotor capture would be expected to be rare in NK-AML cases such as our study cohort, although the translocation of a single gene would not be detectable with standard karyotyping techniques.
A large SNP association study, based on data from The Cancer Genome Atlas and the Ovarian Cancer Association Consortium on 4616 chemotherapy treated ovarian cancer patients, indicated that the 1236C>T, 2677G>T/A and 3435C>T SNPs were the best candidates as prognostic markers, but only a small influence of the 1236C>T on survival could further be demonstrated in data from the 14 studies included (114). This study can be considered to some extent as supporting our results in AML, albeit toning down the magnitude of the effect.

Two more recent meta-analyses also report associations between \textit{ABCB1} SNPs and overall survival in AML with standard induction treatment (49, 115). Our Paper I and Paper III were included as part of the meta-analyses (a total of seven publications met the inclusion criteria), and concluded that there was a positive influence of all the three above-mentioned SNPs and OS as well as CR. Our studies do not support any correlation for the 3435C>T variant and we did not find any association with the chance of CR, but the meta-analyses largely demonstrate that despite the inconclusiveness of previous studies in different settings, \textit{ABCB1} SNPs should not be disregarded as potential markers of prognosis after standard induction treatment of NK-AML. Such prognostic markers would be of particular value in the group of patients with double-negative \textit{FLT3}-ITD/NPM1, where the outcomes have high variability and other risk-stratifying molecular markers are missing. However, more prospective and/or functional studies, incorporating both pharmacokinetics and investigating transport-independent functions of ABCB1 in relation to different genotypes, need to be performed before any incorporation of \textit{ABCB1} genotyping into clinical practice can be recommended.
SNPs in genes related to Ara-C metabolism: Utility as markers of prognosis

The three genes CDA, DCK and cN-II were investigated in a cohort of 205 NK-AML patients based on the hypothesis that SNPs affecting the function/activity of the gene products could influence the relative balance of drug activation and inactivation, leading to variations in intracellular levels of the active metabolite Ara-CTP and thereby influencing patient outcome (64). Outcome measures included chance of CR, overall survival and progression-free survival.

Of the SNPs investigated, only the DCK 3'-UTR *165C>T (rs4643786) variant showed any indication of an association with the chance of CR, with all 16 heterozygotes achieving CR compared to 82% of the homozygous T/T (p=0.08). SNPs in the 3'-UTR region of DCK have previously been related to DCK mRNA expression and levels of Ara-CTP in AML blasts (70), which could explain differences in the chance of a good treatment response. However, in our case the differences could also be explained by the younger age of the heterozygous patients, who were also more frequently transplanted.

Relationships with overall survival

In terms of survival measurements, the SNPs did not appear to have any significant influence when the entire material was analysed in univariate survival analysis. In a second step, like that in Paper III, the material was stratified based on FLT3 status. This indicated an effect on overall survival from the two CDA variants -451C>T (rs532545) and 79A>C (rs2072671) in patients with simultaneous FLT3-ITD (Figure 8). Since these two SNPs were strongly linked, the haplotype was also considered in the multivariable analysis (comparing carriers of the homozygous CC/TT to all others), together with age, gender, FLT3-ITD/NPM1-group and treatment (chemotherapy alone or chemotherapy followed by allo-SCT). Unstratified Cox regression analysis demonstrated a significant negative influence of age, FLT3-ITD+/NPM1- and CDA 79C/C and/or -451T/T genotype on survival. FLT3-ITD+/NPM1+ genotype also appear to confer a slightly worse prognosis (compared to the FLT3-ITD-+/NPM1- set as the reference), and allo-SCT was associated with improved outcomes for all groups. Stratified Cox regression supported the indication that the effect of the CDA genotype was most pronounced in the FLT3-ITD+ subgroup.
Figure 8. OS for CDA genetic variants, with significant differences in patients with FLT3-ITD. Mean OS was 5 vs. 23 months and 5 vs. 24 months, for A and B respectively.
Both CDA SNPs studied have previously been demonstrated to correlate with lower mRNA expression for the alternative alleles compared to the wild type (72, 75), and functional in vitro studies confirm reduced activity for enzymes with the 27Gln residue corresponding to the 79A allele (68, 116).

Decreased deamination of Ara-C, giving higher intracellular concentration of the active metabolite could potentially increase the drug's efficacy. However, higher intracellular concentrations could also mean increased risk of severe toxicities. This has been demonstrated in an AML study by higher incidence of Ara-C-related severe liver toxicity in patients with the CDA 79C allele, and shorter survival in those with CDA -451T/T genotype (74). A more recent article on a mixed AML population also showed correlations between the variant 79C allele and survival times, and increased risk of severe skin toxicities with the -451T variant (76), although here the poor outcome related to the 79A>C variant was significant in the heterozygous A/C patients (no hazard ratio reported for the homozygous C/C), and the authors noted that the genotype frequencies for both those SNPs was out of Hardy-Weinberg equilibrium in their population. However, both studies support an association between variant alleles of these two CDA SNPs and a worse treatment outcome in AML and are in accordance with our results, although no toxicity data were available for our cohort.

In addition to the impact on survival shown for CDA SNPs, there was also some indication that the cN-II 7A>G (rs10883841) had an influence on survival, but this was not significant and inclusion of both FLT3 and NPM1 status reduced the differences. Differences in mRNA expression in relation to genetic variation of this gene have been demonstrated earlier (72). Other SNPs may still prove to have a prognostic value, but in conclusion our results do not support any clinically relevant potential of the rs10883841 variant as a marker of prognosis in NK-AML.

More studies in large, well-defined AML cohorts are needed, but CDA 79A>C and -451C>T as markers of prognosis could be of high value especially in the FLT3-ITD+/NPM1+ cases, to further distinguish relatively high-risk patients within this heterogeneous group, who may benefit from early allo-SCT. Potential relationships between CDA genotype and increased risk of severe toxicities need to be investigated deeper in AML during standard treatment, before any certain conclusions can be drawn regarding optimal preventive actions or dose adjustments based on genotype.
**In vitro data**

Somewhat paradoxically, the *in vitro* data from a small subset of the patients showed an increased sensitivity for the CDA homozygous variants towards topoisomerase inhibitors but not towards Ara-C. The lack of difference in Ara-C sensitivity could be explained by methodological difficulties including assay variability or because, at concentrations representative of the *in vivo* situation, the cells are mainly kept alive in the short-term *in vitro* incubations without extensive cell division, and therefore any actual differences between CDA SNP variants may not be detectable.

The increased sensitivity towards topoisomerase inhibitors in relation to CDA genotype is more difficult to explain. One possible explanation for this result could be that the homozygous 79C/C and -451T/T patients were also carrying the *ABCB1* 1236T/T and 2677T/T genotypes, which were associated with increased sensitivity. Since the topoisomerase inhibitors are substrates of ABCB1 but not of CDA, differences would more likely be attributable to the transporter than to the CDA enzyme activity. However, this was not the case; only a fraction of the few CDA homozygote variants in the *in vitro* subset were also homozygotes for the *ABCB1* variants (12.5% of the cases were homozygote 1236T/T and 16.7% of the cases were homozygous 2677T/T, unpublished data).

A small experiment on 82 of the patients indicated a decreased global DNA methylation related to the variant alleles, although the number of rare homozygotes was low. This raised the question of whether the difference in topoisomerase inhibitor sensitivity could be related to the methylation status, although it is not clear how or if the CDA enzyme is related to DNA methylation.

Another hypothesis is that altered CDA activity due to the SNPs may lead to altered expression of topoisomerase enzymes, thereby influencing the drug effect. This has been demonstrated for another member of the family, activation-induced cytidine deaminase, which has been shown to suppress topoisomerase 1 mRNA translation through deamination (117). So far there have been no publications of a similar function for CDA, and unfortunately topoisomerase RNA levels could not be investigated in our patients.
**Prognostic value of IDH2 mutations and IDH1 codon 105 SNP**

The IDH1 and IDH2 genes have been found to be recurrently mutated in AML, and the mutations lead to neomorphic enzyme activity with generation of the oncometabolite 2-HG instead of the normal α-KG. Mutations have been associated with impaired differentiation and hypermethylation, and we investigated the impact of IDH mutations as well as the IDH1 codon 105C>T SNP (rs11554137) on prognosis in an unselected cohort of 189 AML patients.

**IDH1 and IDH2 mutations**

Baseline characteristics did not differ significantly between the separate IDH genotypes, but combined data indicated that patients carrying mutation in either gene were slightly older at diagnosis compared to those without any IDH1/2 mutation. Previous studies support the finding of older age in association with IDH mutations, but also indicate that patients with IDH mutations have other clinical features in common such as higher platelet count, higher bone marrow and peripheral blood blast counts, and a higher degree of neutropenia at diagnosis (105, 118, 119). We did not have data on bone marrow and blood counts at time of diagnosis and could not therefore investigate such associations.

In terms of induction treatment response, we could not demonstrate any impact of IDH1/2 mutations or genotype of the IDH1 codon 105 SNP. No significant influence of the IDH1 codon 132 mutation on OS was found, either in the entire cohort or in the different risk groups. Multivariable analysis showed that IDH2 codon 140 mutated patients displayed a significantly shorter OS compared to unmutated patients, with age, risk classification and treatment (chemotherapy only or chemotherapy + allo-SCT) as additional significant factors for survival (Table V). The effect on survival appeared to be most pronounced in the intermediate risk group.

In contrast, patients with the IDH2 codon 172 mutation had an improved survival compared to codon 172 wild type patients, which is in accordance with the more recent finding by Papaemmanuil et al. that IDH2 R172 mutated patients have a long-term survival comparable to that of NPM1 mutated patients (5). However, in both our study and theirs, the number of patients with this mutation was limited. A few studies have indicated that the effect of IDH2 mutations may be dependent on the location of the mutation. Marcucci et al. (86) reported lower CR rate for the IDH2 R172 mutation compared to IDH1/2 wildtype, while three other studies demonstrated improved OS associated with the IDH2 R140 mutation (81, 106, 120), which is not in agreement with our results. In
contrast to our results, Green et al (106) also found an impaired outcome for patients with the \textit{IDH2} R172 mutation. There are some differences between our cohorts and theirs, such as higher median age in our patients compared to mainly younger patients in the studies by Green et al. (106) and Patel et al. (120).

However, it is still not clear why there would be such divergent effects on outcome due the \textit{IDH2} R140 and R172 mutations since all \textit{IDH} mutations identified have been shown to result in increased 2-HG levels, with \textit{IDH2} R172 appearing to have slightly higher levels compared to the \textit{IDH1} R132 and \textit{IDH2} R140 (81, 82, 121). Some differences were noted by Papaemmanuil et al., such as the association with \textit{NPM1} mutation for the \textit{IDH2} R140 mutation but mutual exclusivity for the \textit{IDH2} R172 mutation with \textit{NPM1} and many other recurrent lesions (5). Differences in gene expression and methylation profiles for the \textit{IDH2} R172 mutation compared to the other \textit{IDH} mutations have also been reported, indicating that the R172 may be related to a more severe phenotype (86, 122). This led to the idea that \textit{IDH2} R172 mutated AML should be viewed as a distinct entity.

Studies prior to ours have not been conclusive in terms of the prognostic value of the \textit{IDH} mutations in AML. Marcucci et al. (86) reported that in their cohort of NK-AML, younger patients with the low risk \textit{NPM1} mutated/\textit{FLT3}-ITD- combination, \textit{IDH1} mutation was associated with an impaired disease-free survival compared to the corresponding \textit{IDH1/2} wild type patients. A lower response rate for \textit{IDH2} R172 mutated cases was also observed. Like ours, two other studies (101, 102) investigating large AML cohorts, including both aberrant and normal karyotypes, failed to demonstrate any significant impact in the entire cohorts but identified relationships between \textit{IDH} mutations and outcome in different patient subsets. Both studies reported that in NK-AML with \textit{NPM1} mutation and without \textit{FLT3}-ITD, \textit{IDH} mutations conferred an impaired prognosis.

In contrast to our results, \textit{IDH1} mutation also had a negative influence on outcome, but in summary these studies confirm that the influence of \textit{IDH} mutations is dependent on patient subset, including age and other co-existing mutations as important variables. Notably, these studies found an association between \textit{IDH1} mutations, intermediate risk or NK-AML and \textit{NPM1} mutation (101, 102). We did not find any such associations, although this could be due to the relatively high degree of missing data for some of these parameters. This included \textit{NPM1} status which was missing in a total of 20% of our patients; 5/15 (33%) with \textit{IDH1} mutation, 2/21 (9.5%) with \textit{IDH2} R140 mutation and 2/5 (40%) with \textit{IDH2} R172 mutation were lacking information on this marker.

A meta-analysis based on 33 selected studies including ours also came to the conclusion that \textit{IDH1} mutations are associated with poor prognosis
while *IDH2* mutations, especially R140, seem to correlate with improved survival, but only in certain groups such as younger patients with intermediate risk AML (123). The authors noted that cases with R172 mutation appear to have adverse outcomes, although the entire meta-analysis was influenced by large heterogeneity among the included studies. Therefore, the differences and exact influence on prognosis for different *IDH* mutations remain somewhat elusive.

Table V. Multivariable analysis of the influence of *IDH1/2* mutations and the *IDH1* codon 105 SNP on overall survival.

<table>
<thead>
<tr>
<th>Covariates</th>
<th>N</th>
<th>HR</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.022</td>
<td>1.002-1.042</td>
<td><strong>0.033</strong></td>
<td></td>
</tr>
<tr>
<td>Risk group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low risk</td>
<td>32</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate risk</td>
<td>87</td>
<td>2.980</td>
<td>1.495-5.942</td>
<td><strong>0.002</strong></td>
</tr>
<tr>
<td>High risk</td>
<td>55</td>
<td>5.993</td>
<td>2.912-12.333</td>
<td>&lt;<strong>0.001</strong></td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>118</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemotherapy + allo-SCT</td>
<td>56</td>
<td>0.231</td>
<td>0.118-0.450</td>
<td><strong>&lt;0.001</strong></td>
</tr>
<tr>
<td><em>IDH1</em> codon 132</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wild type</td>
<td>162</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mutated</td>
<td>12</td>
<td>0.816</td>
<td>0.390-1.708</td>
<td>0.59</td>
</tr>
<tr>
<td><em>IDH2</em> codon 140</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wild type</td>
<td>155</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mutated</td>
<td>19</td>
<td>1.942</td>
<td>1.068-3.530</td>
<td><strong>0.030</strong></td>
</tr>
<tr>
<td><em>IDH2</em> codon 172</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wild type</td>
<td>169</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mutated</td>
<td>5</td>
<td>0.222</td>
<td>0.067-0.738</td>
<td><strong>0.014</strong></td>
</tr>
<tr>
<td><em>IDH1</em> SNP codon 105 GGC&gt;GGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wild type</td>
<td>154</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variant</td>
<td>20</td>
<td>1.496</td>
<td>0.812-2.756</td>
<td>0.196</td>
</tr>
</tbody>
</table>

95% CI=95% confidence interval; HR=hazard ratio
**IDH1 codon 105 SNP**

The IDH1 codon 105 GGC>GGT SNP (rs11554137) is a synonymous variant located in the same exon as the R132 mutation and was first reported by Wagner et al. (88) as a marker of prognosis in NK-AML. Their study included 275 NK-AML patients of relatively young age – 17-60 years, median 47 years – and they also analysed the IDH1 R132 mutation. While the R132 had a trend towards lower response rates, no association with survival was found in the entire cohort or in subgroups based on NPM1/FLT3-ITD status. However, an inferior survival was seen for the variant allele of the SNP. The authors also reported an increased mRNA expression associated with the SNP T allele, in contrast to the decreased expression for the IDH1 R132 mutation.

In our cohort, 20 patients (10.6%) with the alternative allele were identified, with only seven of those cases having an overlap with IDH mutations. Survival was inferior in patients carrying the alternative allele, and although this has a general concordance with the results of Wagner et al., they found that the impact on survival was seen mainly in patients with either NPM1 wildtype or FLT3-ITD while our subgroup analysis indicated that the effect of the SNP was limited to the intermediate risk patients without FLT3-ITD (Figure 9). It should be noted that among these FLT3-ITD negative intermediate risk patients with the alternative SNP allele, all were NPM1 wild type. This indicates that the codon 105 variant could have potential as a clinically important tool for distinguishing relatively high-risk patients among the heterogeneous intermediate risk patients lacking prognostically relevant mutations in FLT3 and NPM1.

Prior to our study and in addition to that carried out by Wagner et al., three other studies have investigated the prognostic implications of the rs11554137 SNP in AML. In two of the studies, SNP genotype did not significantly influence outcome in childhood AML (90, 91), whereas one of those publications demonstrated an inferior EFS in adult patients with the alternative T allele (91). The third article failed to demonstrate any influence of the SNP (or IDH mutations) on outcome in their cohort of patients; notably, all were treated with different protocols including high-dose Ara-C and the newer agents tipifarnib (a RAS kinase inhibitor), sorafenib (a tyrosine kinase inhibitor) or vorinostat (a histone deacetylase inhibitor) (124). Consequently, it is possible that dose adjustments or co-administration of more targeted treatments may overcome the influence of IDH aberrations, explaining the lack of significant correlations in this study.

Correspondence published a short time before our study also reported that increased levels of 2-HG in AML patients were associated with the SNP, in the absence of co-occurring IDH1/2 mutations (125), providing further evidence of a biological significance. The meta-analyses mentioned
above also support the prognostic impact of the SNP in AML, and for those analyses no significant heterogeneity among the included studies was apparent (123).

Figure 9. Influence of IDH codon 105 SNP rs11554137 on overall survival in intermediate risk FLT3-ITD negative patients. Median OS 6 months vs 20 months.
Risk-adapted therapy based on *IDH* status

As a commonly mutated gene in AML, *IDH* status is likely to become important in adaption of the standard treatment protocols, both as a biomarker for patient selection to new targeted treatments and as a prognostic marker that may indicate subgroups of patients eligible for allo-SCT. While early allo-SCT is a given option for patients classified as high-risk according to the current guidelines, more tools are needed to distinguish relatively high-risk patients among the heterogeneous intermediate risk group.

Based on our data, intermediate risk AML patients lacking further prognostic guidance from *FLT3*-ITD or *NPM1* mutational status and carrying an *IDH2* mutation or the *IDH1* codon 105 SNP variant allele could be further assessed as candidates for early allo-SCT. Other factors such as age, performance status and co-morbidities should, as always, be considered together with mutation/SNP genotype.

As previously mentioned, high-dose chemotherapy could potentially overcome some of the effects of the mutations. However, given the higher age of patients with *IDH* mutations, such dose escalations may not always be tolerated, and other options should be considered. The hypermethylated DNA profile demonstrated for *IDH* mutant haematopoietic cells is associated with impaired differentiation through the impairment of *TET2* mediated demethylation (84). This implicates the use of hypomethylating agents such as azacytidine or decitabine, in an attempt to normalize the epigenetic status and favour cell differentiation. While some clinical trials indicate an improved outcome in *IDH* mutated patients with these drugs, others have not found the same (as reviewed by Montalban-Bravo & DiNardo 2018) (126). Whether this is due to differences in patient cohorts, sample size etc. is not clear, and more studies are needed as newer hypermethylating agents are being developed.

New treatments targeting the neomorphic *IDH* mutant enzyme have reached clinical trial and in 2017 the first *IDH2* small molecule inhibitor enasidenib was approved by the FDA. The drug binds to the active site of the mutated enzyme and inhibits the conversion of α-KG to 2-HG, leading to subsequent differentiation of the AML blasts (127). This and other small molecule inhibitors are currently in trial in different populations of *IDH1/2* mutated haematological malignancies. Like the treatment of APL where a combination of ATRA and arsenic trioxide is used, it is not unlikely that the efficacy of *IDH* inhibitors (which induces differentiation but not apoptosis to any higher degree) could be improved by combination treatment with classic chemotherapeutics.

Our data and others also implicate functional and prognostic effects for the constitutional *IDH1* codon 105 SNP similar to that of *IDH* mutations,
and it should be further investigated whether patients with the alternative SNP allele would benefit from the new targeted treatments. The primary candidates who could be suggested for enrollment are older, intermediate risk patients without FLT3-ITD and NPM1 mutations, and other factors that speak against the use of standard intensive chemotherapy.
CONCLUSIONS

In conclusion, results from Papers I-IV indicate that genetic variants in *ABCB1*, *CDA* and *IDH1* as well as *IDH2* mutations have potential as prognostic markers in acute myeloid leukemia, that they have implications for drug sensitivity, and that they could be useful for better risk assessment and guiding treatment decisions. The most promising of the different variants analysed are as follows:

- Papers I and III: *ABCB1* SNP 1236C>T (rs1128503) and 2677G>T (rs20325829) in NK-AML, with higher *in vitro* drug sensitivity in patient cells and increased OS primarily in FLT3-ITD negative cases carrying the alternative alleles;
- Paper II: *CDA* SNP 79A>C (rs2072671) and -451C>T (rs532545) in NK-AML, with inferior prognosis in FLT3-ITD positive cases with the homozygous C/C and T/T genotypes;
- Paper IV: *IDH2* R140, with shorter OS in mutated cases, and *IDH2* R172, with indications of improved OS in mutated cases. In the subset of intermediate risk FLT3-ITD negative AML patients (including both normal karyotype and patients with karyotypes of unclear significance), the *IDH1* codon 105 SNP (rs11554137) variant allele is associated with decreased survival times.
It can be concluded that the influence of the different variants depends on other factors, primarily cytogenetic risk and FLT3-ITD. Given the emergence of IDH targeting therapies, IDH1/2 mutations and the IDH1 SNP are likely to be the markers of this thesis with the nearest future of incorporation into treatment guidelines and clinical practice. A graphical presentation of the different prognostic subgroups indicated by the results from Papers I-IV is outlined below (Figure 10).
Figure 10. Prognostic subgroups as indicated by the results of Papers I-IV.

Always consider other leukemia- and patient-specific risk factors!
Prognostic markers in acute myeloid leukemia – A candidate gene approach
Despite the positive results of this thesis on the impact of ABCB1 SNPs on the outcome in AML, the collective data published so far on the relevance of genetic variation of this gene in different settings is inconclusive. In a screening setting, a combination of both genotyping and phenotyping could provide more information on the clinical relevance of ABCB1 for the patient in terms of risk classification. The same would be true for the CDA variants, where information on toxicity should also be included in any future evaluations.

In all clinical trials of new and older drugs, pharmacogenetic evaluations could be integrated in the study design, with well-selected gene panels or whole genome/exome analysis. In addition to this, pharmacokinetic data including plasma and urine concentrations of the parental drug as well as metabolites are of high value. For both the IDH mutations and the IDH1 codon 105 SNP, enantiomer-specific analysis of 2-HG together with methylation analysis and RNA expression profiles could add additional information on the biological consequences of these different alterations.

Given the heterogeneity of cancer cells in AML as well as other tumour types, research combining new technologies such as single cell or bulk RNA sequencing with genotyping may further elucidate the role of different SNPs in both normal cells and tumour cells of different molecular
backgrounds. We currently have a study underway where sequential samples taken from AML patients before and during the first day of induction treatment have been analysed by RNA sequencing. The pre-treatment sample will be used in the upcoming bioinformatic analyses as a reference to identify genes up- or down-regulated in response to the treatment. This exploration of the earliest changes in expression will hopefully give new insights into which single genes, complex networks of genes, or pathways that are important for chemotherapy response and resistance. Preferably, relapse samples should also be included in follow-up analysis, and long-term outcome will be followed with the hope of identifying whether certain early changes in gene expression levels could be associated with end-points such as overall survival and thereby be of use as markers for better risk classifications.

Figure 11. Future NGS-based studies of AML in relation to treatment response and outcome.

In addition, deep sequencing of sequential samples could help in distinguishing sensitive and resistant clones. Mutations that are harbouried by the more resistant clones would be expected to remain or increase in relative proportion, while others shared by cells from the more sensitive clone are eradicated by standard chemotherapy. Later, it would be those
resistant clones that re-emerge at relapse. This approach could help identify mutations, or combinations of mutations, associated with resistance and impaired prognosis. As more and more targeted drugs are developed, this knowledge would help in guiding different patients to the most optimal treatment strategy based on individual mutational patterns.

Voices have also been raised to push the field of personalized medicine beyond pure genomics, with the integration of more functional approaches including 3D cultures of primary tumour cells and patient-derived xenografts (128). This is certainly needed in order to elucidate the functional consequences of cancer-related mutations as well as for the discovery of other vulnerabilities that can be targeted in cancer treatment. As for clinical integration of functional approaches for acute leukemias, there are some unfortunate limitations. The most important one is the timeline; it is simply not possible to wait several weeks or months for experimental results before making treatment-related decisions. The idea of taking a simple blood sample and, within a few days, analysing a set of biomarkers to assess risk and predict which treatments that the patient is likely to respond to is certainly tempting, and relative to functional approaches it is also (at least in theory) easier to implement into clinical practice. However, despite the hopes and potentials of pharmacogenomics and other omics-based precision medicine studies, there is still a long way to go before we are close to true individualization of the treatment of acute myeloid leukemia.

The introduction of NGS techniques into clinical routine is currently underway at most larger hospitals in Sweden. The focus has generally been on cancer-specific panels that aid in the diagnosis and classification of cancer type. However, our results show that constitutional genetic variation should not be overlooked, and that continued research is needed in close collaboration with the health care units to decide if and how the analysis of normal genetic variation affecting drug response should be incorporated into the use of these new technologies. Some have suggested that the future holds whole genome sequencing of all individuals even before the debut of any illness, so that the genetic information can be stored and made available without delay. As the scientific evidence for specific genetic markers emerges, selected parts of the data can be accessed when needed for the care of the individual. This places high demands on well-developed bioinformatic systems, regulations and IT security to continuously protect personal integrity while storing and using this type of sensitive information.
ACKNOWLEDGEMENTS

Det finns ett stort antal människor utan vars hjälp denna avhandling aldrig hade kunnat komma till. Så många som hjälpt till att bära mig ända hit, att listan är närmast ofattbar. Familj, vänner, kollegor. Jag har inte glömt någon av er. Ett säregnt tack vill jag rikta till följande, utan inbördes ordning:


Mina närmaste doktorandkollegor Anna Svedberg och Niclas Björn. När arbetet fått mig att slita mitt hår har det alltid ändå känts värt att komma hit, tack vare er. Ni har också fått mig att förbättra mina manusgranskarskills, med uppmuntran genom chokladmutor. Lucia Pelle, our super-Italian post-doc. For your amazing energy and warm personality, wonderful dinners and long talks. I really appreciate all your support.

"Gamla gänget" på Klinfarm. Karin Skoglund, tack för alla djupa samtal, skräckfilmskvällar, spelkvällar, boksnack och road tripping över slätterna. Tack för inspiration till avhandlingsomslaget, för planeringshjälp och

**Britt Sigfridsson**, **Monica Häger** och **Curt Peterson** som fick mig att känna mig välkommen redan som student. **Elisabeth Hamrin**, min första kontorsroomie som såg min doktorsavhandling framför sig redan innan jag tagit min magisterexamen. **Sally Coulthard**, för having confidence in me as an assistance in your projects and teaching me almost everything I know about cell culturing. **Malin Lindqvist Appell**, som anförtrott mig både Pyrolab och studenter, och påmint mig om att det är okej att säga nej ibland. **Susanne Andersson**, min kära MedBi-mamma som alltid hejar på och värnar om oss gamla studenter.

**Madeleine Örlin**, du är vår avdelnings klippa och jag skulle inte ha klarat mig utan det fantastiska jobb du gör, både med administrativ support, för avdelningens allmänna trivsel och för din omtänksamhet när allt kändes hopplöst. Alla andra fina människor på **LÄFO/KVM** som på olika sätt förvann min tillvaro både inom och utanför arbetet genom åren. Särskilt tack till **Simon Jönsson** som följt mig hela vägen från MedBi-nolla till doktorand, blandade mina drinkar, och utan att blinka ställde upp som flytthjälp i en påfrestande situation. **Daniel Söderberg** and **Emina Vorcapić**, för party feeling och många skratt. **Anna Zimdahl Kahlin**, för bra samarbete med studentlab och fint sällskap på kurs i Uppsala. Also a big thank you to all the other awesome Ph.D students (old ones and new), teacher colleagues and researchers who contribute joy and inspiration on a daily basis.


Tack till er vänner jag funnit i Internets mörka hörn. **Roland**, för ovärderligt stöd i vardagens vedermödor, teknisk support, för att du alltid lyssnar, ger mig perspektiv och peppar mig att fortsätta min personliga utveckling. Tack också för lån av katt! Du är en sann vän. **Per**, för att du spontant ryckte ut med chiligryta i en av mina mökare stunder, och för en


Tack till **Linköpings Universitet** som gav mig en plats att utvecklas på. Tack till följande anslagsgivare: **Vetenskapsrådet**, **Cancerfonden**, **AFA försäkrings**, **Region Östergötland Cancerföreningen i Stockholm och Karolinska Institutet**. Tack också till **LiU Cancer** och **Anna Cederbergs stiftelse för medicinsk forskning**, som bidragit till att jag fått åka på stora internationella konferenser och lära mig mer, och till **Lions forskningsfond mot folksjukdomar** som förerat mig doktorandstipendium.
REFERENCES


50. Woodahl EL, Crouthamel MH, Bui T, Shen DD, Ho RJ. MDR1 (ABCB1) G1199A (Ser400Asn) polymorphism alters transepithelial
87. Chotirat S, Thongnoppakhun W, Promsuwicha O, Boonthimat C, Auewarakul CU. Molecular alterations of isocitrate dehydrogenase 1 and 2 (IDH1 and IDH2) metabolic genes and additional genetic mutations


Papers

The papers associated with this thesis have been removed for copyright reasons. For more details about these see:

http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-153476