Pharmacogenetics of osteosarcoma treatment

Hanneke Vos
Pharmacogenetics of osteosarcoma treatment

Colofon

About the cover: On a background of microarray signals, the front cover shows a few branches of the tree shown in full on the back cover, inspired by the story 'Leaf by Niggle' (J.R.R. Tolkien). The tree represents the ultimate aim that we envision to achieve in our work. In the present thesis, this is improved treatment of patients with osteosarcoma. The story conveys the message that the few leaves or branches that we may finalize in our daily work all contribute to the ultimate aim and that there is hope that this aim will be fully reached one day.

The research project was funded by ZonMw.

Printing of this thesis was financially supported by the Radboud university medical center, Nijmegen, and Department of Clinical Pharmacy & Toxicology, Leiden University Medical Center, Leiden, The Netherlands.


Cover design: Proefschrift-AIO
Lay-out: Proefschrift-AIO
Printing: DPP

Copyright © 2016 Hanneke Vos, Utrecht, The Netherlands
Chapter 1 General introduction

METHODS

Chapter 2 High-quality genotyping data from formalin-fixed, paraffin-embedded tissue on the drug metabolizing enzymes and transporters plus array

ADVERSE DRUG REACTIONS

Chapter 3 Replication of a genetic variant in ACYP2 associated with cisplatin-induced hearing loss in patients with osteosarcoma

Chapter 4 The role of genetic variants in ADME genes in high-dose methotrexate-induced hepatotoxicity in patients with osteosarcoma

DRUG EFFICACY

Chapter 5 Pharmacogenetic risk score predictive of treatment response in patients with osteosarcoma

Chapter 6 Genetic variants of drug transporters influence response to chemotherapy-based treatment in osteosarcoma

Chapter 7 Pharmacogenetics of chemotherapy response in osteosarcoma: a genetic variant in SLC7A8 is associated with progressive disease

Chapter 8 General discussion

Summary | Samenvatting

PhD Portfolio
Curriculum Vitae
List of publications
Dankwoord
Chapter 1

General introduction

Adapted from:
The role of pharmacogenetics in the treatment of osteosarcoma
Hanneke I. Vos, Marieke J.H. Coenen, Henk-Jan Guchelaar, D. Maroeska W.M. te Loo

Drug Discovery Today 2016, in press
Introduction

Cancer is one of the major causes of death worldwide. Treatment of cancer varies among the different types of tumors, but classically involves either chemotherapy, radiotherapy, surgery, or a combination thereof. The increasing understanding of tumor biology and advances in the field of genomics have enabled the development of more precise treatment modalities such as immunotherapy and targeted therapy using monoclonal antibodies for specific subsets of tumors, based on predictive biomarkers. In addition, fine-tuning of chemotherapeutic treatment on the basis of insights in the pharmacokinetics and underlying genetics has become possible. All of these efforts, which are forms of personalized medicine, are aimed not only at increasing the patient’s chance to survive, but also at reducing the risk of side effects that can lead to morbidity and mortality in itself, and that have impact on the quality of life of surviving patients. This is especially important for cancers occurring during childhood, such as osteosarcoma, which is focused on in the current thesis.

Osteosarcoma

Osteosarcoma is a primary bone malignancy predominantly affecting children, adolescents and young adults. Although rare with a worldwide incidence of 3-4 per million, it is the most common type of primary bone cancer [1]. The primary site is most frequently the distal femur or proximal tibia and at diagnosis already 20% of patients present with clinically detectable metastases, mostly in the lungs [2,3]. The outcome for patients with high-grade osteosarcoma was poor before the introduction of effective chemotherapy in the late 1970s. Since the addition of multiagent chemotherapy to the treatment regimen, 5-year survival increased from below 20% with surgery alone to 60-65% with combined surgery and chemotherapy [4]. However, the increase in survival has stagnated and still patients relapse, despite intensive treatment. Moreover, survivors often experience long-term chemotherapy-induced toxicities. Hence, there is a significant unmet need to improve survival of patients with osteosarcoma.

The three-drug backbone of the treatment consists of cisplatin, doxorubicin and methotrexate (MTX) (MAP). Although these chemotherapeutics have been found to be the most effective treatment for osteosarcoma in clinical trials during the past decades [5], the sensitivity of patients to these agents, with regard to the antitumor effect and to the toxic side effects, varies highly among patients with osteosarcoma. As is the case for all chemotherapeutics, these agents have a narrow therapeutic window, which poses a challenge to treatment optimization. Further treatment intensification is generally limited due to treatment-related toxicity and mortality.

Currently, the presence of metastases at diagnosis is the only consistent clinical prognostic factor for risk stratification [2]. Nevertheless, the treatment regimen is still similar for all patients regardless of metastatic status at diagnosis. The histologic response (HR) after preoperative chemotherapy has been studied as an approach to distinguish poor responders from good responders at surgery. However, there are no options for substantial changes to the treatment other than extra salvage treatment for patients with a poor response. The recent European and American Osteosarcoma Study Group (EURAMOS)-1 trial, in which the addition of the chemotherapeutics ifosfamide and etoposide (IE) to the adjuvant MAP treatment (resulting in MAPIE) for poor histologic responders was studied, could not provide evidence of an advantage of MAPIE over MAP in event-free survival (EFS) and overall survival (OS), and MAPIE treatment was associated with greater toxicity [6]. Moreover, the histologic response is not a consistent and reliable predictor of outcome in past trials, and is not available at diagnosis but only after the first cycles of standard neoadjuvant chemotherapy [7]. Other than clinical factors, biological biomarkers could also be possible candidates for stratification and development of targeted therapy or treatment intensification, for example ATP-binding cassette sub-family B member 1 (ABCB1) overexpression [8], which is investigated as the basis for intensified treatment using the immunomodulator mifamurtide in the ongoing Italian Sarcoma Group (ISG)/OS-2 trial. However, in routine clinical practice, there are currently no treatment options other than chemotherapy available. Alternative treatment strategies, specifically targeting for example apoptosis resistance, proliferation, angiogenesis and immune evasion, are in development or clinical testing phase [9-12]. This includes agents in combination with chemotherapy such as monoclonal antibodies and interferon alpha 2b. The latter was tested in the EURAMOS-1 trial as maintenance therapy for patients with a good response to chemotherapy; however, the first report does not show an improvement in EFS when compared to chemotherapy alone [13]. Efforts to influence the bone microenvironment using bisphosphonates such as zoledronic acid are still in the early phase of clinical testing [14]. Furthermore, angiogenesis inhibitor sorafenib in combination with mTor inhibitor everolimus appeared to be favorable in a recent Phase II trial [15]. However, further development and testing of these strategies will still take considerable time.

Despite advances in the field of tumor genomics in general, the complex karyotype and inter- and intra-tumor genetic heterogeneity of osteosarcoma is a complicating factor to design treatment regimens directed to the tumor genome. To date, tumor genomics approaches have not resulted in an established (somatic) genetic tumor profile that could already be clinically used as predictor of treatment efficacy [16]. Nevertheless, interesting results have been provided for example by genomic copy number variant studies, pointing towards chromosomal alterations as predictors of clinical outcome, which warrants further validation [17]. However, not only the tumor but also the germline genome is...
Pharmacogenetics of cancer treatment

The response to chemotherapeutic drugs used in cancer treatment is generally determined by a combination of clinical and genetic factors. Pharmacogenetics focuses on the inherited variability of drug response with regard to efficacy and toxicity. In osteosarcoma, efficacy is mostly investigated as survival (disease-free, progression-free, event-free) and HR, and toxicity studies are mostly focused on ototoxicity and cardiotoxicity. Pharmacogenetics involves genetic variation in pharmacokinetics (PK), including transporters and drug metabolizing enzymes. These enzymes can be categorized in Phase I, responsible for reactions such as oxidation, reduction, and hydrolysis, and Phase II, responsible for conjugation reactions such as glucuronidation and glutathione conjugation. Enzymes of the CYP superfamily, glutathione-S-transferases (GST) and uridine diphosphoglucuronosyltransferases (UGT) are known to catalyze the metabolism of the chemotherapeutics used in the treatment of osteosarcoma. Furthermore, most of the drug transport is covered by two types of transport superfamilies, namely ATP-binding cassette (ABC) efflux pumps and solute-linked carrier (SLC) influx transporters. Genetic variants in genes encoding these proteins can contribute to variation in drug levels that reach the target DNA due to reduced uptake and/or increased efflux, which on its turn can influence treatment efficacy and toxicity. On the other hand, pharmacogenetics also involves genetic differences in pharmacodynamics (PD), such as DNA repair mechanisms. Several DNA repair pathways, such as the nucleotide excision repair (NER) pathway, can protect from the deleterious effect of oxidative DNA damage. Variation in genes encoding such proteins can result in enhanced DNA repair and failure of apoptotic pathways induced by the chemotherapeutics, which also impacts efficacy. Different methods can be employed to explore pharmacogenetic markers, including candidate gene analysis, pathways-based gene analysis, and genome-wide association studies (GWAS). To date, the genome-wide method has been a challenge in osteosarcoma because of the need of large patient cohorts. Hence, almost all past pharmacogenetic studies in osteosarcoma are hypothesis-driven (candidate gene or pathway-based) to varying extent. In the few GWASes that included osteosarcoma patients, these patients were only a minority of the total cohort; no pharmacogenetic GWAS has been carried out with only osteosarcoma patients. The findings of pharmacogenetic studies of PK- and PD-related genes in osteosarcoma with regard to drug efficacy and toxicity are described hereafter. All discussed research involved germline pharmacogenetics, unless indicated otherwise.

Pharmacogenetics of drug efficacy

Pharmacogenetic studies in osteosarcoma have focused mainly on candidate genes related to drug transport, DNA repair and detoxification mechanisms, investigating the treatment response by analyzing the HR and survival. An overview of the published studies is shown in Table 1.

**Transporters**

Multiple studies have investigated the role of genetic variation in drug transporters in the efficacy of chemotherapeutic drugs used to treat osteosarcoma patients. All studies have selected candidate genes encoding transporters with a known function in the pharmacology of one or more of the core agents in the treatment regimen, which could impact efficacy of these agents through differences in PK. The first study that has assessed transporters of drugs in the MAP regimen was published by Caronia et al. [18]. The authors have retrospectively analyzed 91 osteosarcoma patients regarding 346 single nucleotide polymorphisms (SNPs) in 24 genes involved in the metabolic pathways of MAP, vincristine and cyclophosphamide. One variant in ATP-binding cassette sub-family C member 3 (ABCC3; rs4148416) and three variants in ABCB1 (rs4148737, rs1128503 and rs10276036 – the latter two in complete linkage disequilibrium) were identified in association with OS and EFS. ABCC3 and ABCB1 encode efflux transporters. ABCB3 is present in the liver, gallbladder, gut and kidney, is a known transporter of MTX, and suggested in cisplatin, doxorubicin, and vincristine transport. ABCB1 encodes P-glycoprotein, is involved in elimination of agents via bile and the kidney and is also known to be overexpressed in many tumors including multidrug-resistant osteosarcoma. P-glycoprotein is a transporter of MTX, doxorubicin and vincristine; although in osteosarcoma cell lines no evidence has been found for its involvement in resistance to MTX. The authors hypothesized that the variants might have an effect on the efflux of the chemotherapeutics and thereby influence the treatment efficacy. However, the functional consequences of the identified variants are yet unclear. Another variant in ABCB1, studied by Xiaohui et al., is also reported in association with a poor HR [19]. Findings similar to the study of Caronia et al. were reported by Liu et al. in a Chinese population of 162 patients with osteosarcoma, demonstrating a poor HR and higher risk of progression or
death in carriers of the C allele of the \textit{ABCB1} rs1128503 variant or the T allele of \textit{ABCC3} rs4148416 [20]. However, conflicting results have been described in two other Chinese studies reporting the effect in the opposite direction for the \textit{ABCB1} variant rs1128503, and a study in which no association with this \textit{ABCB1} variant was detected at all [21-23]. A meta-analysis of the results of the abovementioned studies of \textit{ABCB1} and \textit{ABCC3} and their association with the response revealed that \textit{ABCC3} rs4148416 and \textit{ABCB1} rs1128503 were associated with response, the latter only in the two included Caucasian studies. However, the definition of response and the interpretation of the reported odds ratios (OR) are unclear from this meta-analysis [24].

A recent study of 62 osteosarcoma patients specifically focusing on MTX-related genes has demonstrated an association of variant rs1051266 in the reduced folate carrier 1 (\textit{RFC1}) gene with survival after MTX-based chemotherapy [25]. Because this was the first reported study investigating this gene in this context, the result warrants replication in independent cohorts, and the true value of the association found is therefore not known.

\textbf{Detoxification}

Several members of the GST family have been studied in relation to the HR and survival of osteosarcoma patients. As the GST enzymes are known to be involved in the detoxification of the predominant chemotherapeutics in osteosarcoma treatment, cisplatin and doxorubicin, but also of agents used in specific protocols such as etoposide, ifosfamide and vincristine, variation in the genes causing differential enzymes activity is likely to be able to influence efficacy of these agents, for example by lowering intracellular drug levels. A total of nine variation in PD, have been investigated by the largest number of studies.

\textbf{DNA repair}

Of all the pathways discussed here, polymorphisms in NER genes involved in the repair of DNA adducts induced by cisplatin, which thereby influence cisplatin efficacy through variation in PD, have been investigated by the largest number of studies. A total of twelve studies have included variants in excision repair cross-complementation group 1 (\textit{ERCC1}) and/or excision repair cross-complementation group 2 (\textit{ERCC2}) [26,30,33-42]. The \textit{ERCC1} variant rs11615 has been reported to be significantly associated only in studies of Chinese patient populations, showing conflicting results. Whereas four studies have suggested a good HR and improved OS for patients with the CC genotype [35,36,40,41], the same genotype was reported in association with inferior EFS or OS by others [37,39]. Of two studies that have also assessed the \textit{ERCC1} rs2298881 variant, only Sun \textit{et al.} have found an association of the AC/CC genotypes with a poor HR and inferior OS [35,40], \textit{ERCC2} variants have also been the subject of inconsistent results. Whereas patients with the rs13181 CC genotype showed improved EFS in an Italian study of 130 osteosarcoma patients, a Spanish study reported an association with a poor HR and inferior EFS, and Chinese study reported inferior OS [33,34,39]. A meta-analysis seeking to determine the effect of this variant has unfortunately only included two of the 12 studies that have investigated this variant. The Italian study has also determined improved EFS of patients carrying the rs1799793 AA genotype, and three other studies have shown similar findings demonstrating improved EFS or OS and a good clinical response [26,37,38]. However, Cao \textit{et al.} have reported an association of this genotype with a poor HR, although no significance was reached in association with OS [40]. There are several differences between the studies, regarding inclusion of metastatic patients or nonmetastatic only, and ethnic background that could have contributed to the discrepancy in the results. Moreover, all studies are reporting on single and relatively small cohorts, often leading to large confidence intervals of the observed point estimates. Follow-up on these results is clearly needed because the studies carried out so far remain inconclusive.


died in carriers of the C allele of the \textit{ABCB1} rs1128503 variant or the T allele of \textit{ABCC3} rs4148416 [20]. However, conflicting results have been described in two other Chinese studies reporting the effect in the opposite direction for the \textit{ABCB1} variant rs1128503, and a study in which no association with this \textit{ABCB1} variant was detected at all [21-23]. A meta-analysis of the results of the abovementioned studies of \textit{ABCB1} and \textit{ABCC3} and their association with the response revealed that \textit{ABCC3} rs4148416 and \textit{ABCB1} rs1128503 were associated with response, the latter only in the two included Caucasian studies. However, the definition of response and the interpretation of the reported odds ratios (OR) are unclear from this meta-analysis [24].

A recent study of 62 osteosarcoma patients specifically focusing on MTX-related genes has demonstrated an association of variant rs1051266 in the reduced folate carrier 1 (\textit{RFC1}) gene with survival after MTX-based chemotherapy [25]. Because this was the first reported study investigating this gene in this context, the result warrants replication in independent cohorts, and the true value of the association found is therefore not known.

\textbf{Detoxification}

Several members of the GST family have been studied in relation to the HR and survival of osteosarcoma patients. As the GST enzymes are known to be involved in the detoxification of the predominant chemotherapeutics in osteosarcoma treatment, cisplatin and doxorubicin, but also of agents used in specific protocols such as etoposide, ifosfamide and vincristine, variation in the genes causing differential enzymes activity is likely to be able to influence efficacy of these agents, for example by lowering intracellular drug levels. A total of nine studies have analyzed the role of GST mu 1 (\textit{GSTM1}) and GST theta 1 (\textit{GSTT1}) deletions [18,20,22,23,26-30]. Only a study of a Brazilian cohort of 80 osteosarcoma patients showed an association of the \textit{GSTM1} null genotype with increased risk of recurrence to the lungs. In addition, metastatic patients with \textit{GSTM1} or \textit{GSTT1} null genotype or homozygous \textit{GSTM1}\textit{A} genotype showed inferior OS [27]. This is the only study that has stratified for metastatic osteosarcoma, which could have contributed to the unique findings. Windsor \textit{et al.} have also detected an association of the \textit{GSTT1} null genotype with inferior progression-free survival (PFS); however, significance was not maintained when corrected for primary site and metastatic disease [23].

Primarily consistent findings have been reported for the GST pi 1 (\textit{GSTP1}) gene. The \textit{GSTP1} rs1695 G allele or GG genotype has been reported in association with a poor HR in multiple studies [20,22,23,29]. Moreover, it has been linked to inferior survival (EFS, PFS, OS, depending on what has been studied) [20,22,23,28,29]. In contrast, one Chinese study that has not investigated the HR but the clinical response instead has identified the GG genotype in association with a good response (complete or partial remission) rather than a poor response; however, this did not translate into inferior OS [30]. A meta-analysis including the total of six studies has confirmed the association of the variant with a poor HR and inferior PFS and OS [31]. Unfortunately, the studies by Yang \textit{et al.} [30] and Zhang \textit{et al.} [28], not showing significance with OS in the meta-analysis, were analyzed separately, because originally the variant rs947894 assessed by these authors appeared to be a distinct SNP, whereas it has been merged with the rs1695 variant later on. Although the findings for rs1695 have been primarily consistent, confirmation in larger study cohorts and clinical validation is needed before it could be translated to clinical practice. Furthermore, another \textit{GSTP1} variant, rs1138272, has been shown to be associated with inferior EFS and OS in tumor DNA [26]. Another meta-analysis has also studied variants in multiple GST genes; however, multiple different phenotypes (toxicity, survival, osteosarcoma susceptibility) were studied together which makes it difficult to draw conclusions [32].
<table>
<thead>
<tr>
<th>No. of osteosarcoma patients (overall)</th>
<th>Ethnic origin; nationality</th>
<th>Drugs column indicates drugs included in treatment of the osteosarcoma patients as far as stated; drugs between brackets are not given to all patients in the study.</th>
<th>Outcome investigated</th>
<th>Genes investigated</th>
<th>Associated variant(s) (Possible) functional role of associated variant</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=91</td>
<td>NS, Spanish</td>
<td>MARS (VCR, CTX, ACD)</td>
<td>HR, EFS</td>
<td>ENRC1, ENRC2, ENRC 4, ENRC5, XPA, APC</td>
<td>ERCC2 rs113811 Controversial&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[33]</td>
</tr>
<tr>
<td>N=130</td>
<td>NS, Italian</td>
<td>MARS</td>
<td>EFS, OS</td>
<td>ENRC1, ENRC2, ENRC 5</td>
<td>ERCC2 rs131811 Controversial&lt;sup&gt;a&lt;/sup&gt;</td>
<td>[34]</td>
</tr>
<tr>
<td>N=40</td>
<td>Caucasian, Slovenian</td>
<td>AMP (MARS)</td>
<td>HR, EFS, OS</td>
<td>ENRC1, ENRC2, GSTM1, GSTP1, GPT1, MBN, MARS, MARS1, APC</td>
<td>ERCC2 rs1799931 DNA repair capacity</td>
<td>[35]</td>
</tr>
<tr>
<td>N=182</td>
<td>NS, Chinese</td>
<td>M, A, P, CTX, VCR</td>
<td>CR, PPS, OS</td>
<td>CON1, BCC1, ENRC1, MARS1, APC</td>
<td>ERCC1 rs1799791 DNA repair capacity</td>
<td>[36]</td>
</tr>
<tr>
<td>N=172</td>
<td>NS, Chinese</td>
<td>MARS</td>
<td>HR, OS</td>
<td>ENRC1, ENRC2, ENRC 4, ENRC5, XPA, APC</td>
<td>ERCC2 rs171051 SEE1 expression</td>
<td>[37]</td>
</tr>
<tr>
<td>N=40</td>
<td>NS, Chinese</td>
<td>MARS</td>
<td>CR, PPS, OS</td>
<td>CON1, BCC1, ENRC1, MARS1, APC</td>
<td>ERCC1 rs2298881 SEE1 expression</td>
<td>[38]</td>
</tr>
<tr>
<td>N=260</td>
<td>NS, Chinese</td>
<td>MARS</td>
<td>CR, PPS, OS</td>
<td>CON1, BCC1, ENRC1, MARS1, APC</td>
<td>ERCC1 rs2298881 SEE1 expression</td>
<td>[39]</td>
</tr>
<tr>
<td>N=186</td>
<td>NS, Chinese</td>
<td>MARS</td>
<td>HR, OS</td>
<td>ENRC1, ENRC2, MARS1, APC</td>
<td>ERCC1 rs171051 See expression</td>
<td>[40]</td>
</tr>
<tr>
<td>N=214</td>
<td>NS, Chinese</td>
<td>MARS</td>
<td>CR, PPS, OS</td>
<td>CON1, BCC1, ENRC1, MARS1, APC</td>
<td>ERCC1 rs171051 See expression</td>
<td>[41]</td>
</tr>
<tr>
<td>N=40</td>
<td>NS, Chinese</td>
<td>MARS</td>
<td>CR, PPS, OS</td>
<td>CON1, BCC1, ENRC1, MARS1, APC</td>
<td>ERCC1 rs171051 See expression</td>
<td>[42]</td>
</tr>
<tr>
<td>N=80</td>
<td>White, Asian, Brown/ Black unclassified, Brazilian</td>
<td>MARS</td>
<td>HR, OS</td>
<td>ENRC1, ENRC2, MBN, A/R, G/R, MARS1</td>
<td>ERCC2 rs1799931 DNA repair capacity</td>
<td>[43]</td>
</tr>
<tr>
<td>N=182</td>
<td>Caucasian, Slovenian</td>
<td>AMP (MARS)</td>
<td>CR, PPS, OS</td>
<td>CON1, BCC1, ENRC1, MARS1, APC</td>
<td>ERCC1 rs1799931 DNA repair capacity</td>
<td>[44]</td>
</tr>
<tr>
<td>N=260</td>
<td>NS, Chinese</td>
<td>MARS</td>
<td>CR, PPS, OS</td>
<td>CON1, BCC1, ENRC1, MARS1, APC</td>
<td>ERCC1 rs2298881 SEE1 expression</td>
<td>[45]</td>
</tr>
<tr>
<td>N=186</td>
<td>NS, Chinese</td>
<td>MARS</td>
<td>CR, PPS, OS</td>
<td>CON1, BCC1, ENRC1, MARS1, APC</td>
<td>ERCC1 rs1799931 DNA repair capacity</td>
<td>[46]</td>
</tr>
</tbody>
</table>

<sup>a</sup> Drugs column indicates drugs included in treatment of the osteosarcoma patients as far as stated; drugs between brackets are not given to all patients in the study.

<sup>b</sup> No germline, but tumor DNA was analyzed.

NS, not specified; HR, histologic response; OS, overall survival; EFS, event-free survival; PPS, progression-free survival; DFS, disease-free survival; CR, clinical response (complete/partial response vs. stable/progressive disease); M, methotrexate; A, adriamycin (doxorubicin); P, cisplatin; I, ifosfamide; E, etoposide; VCR, vincristine; CTX, cyclophosphamide; ACD, actinomycin D; C, carboplatin.
Other DNA repair genes for which significant associations have been reported include the xeroderma pigmentosum complementation group C (XPC) gene, showing a good HR but no improvement of EFS with the rs2229001 AC/CC genotype [33]. Furthermore, ERCC5 rs1047768 TT has been associated with a good clinical response and improved PFS and OS [43], and the rs2296147 variant with good clinical response and improved OS [44,45]. For MMS19 homolog, cytosolic iron-sulfur assembly component (MMS19L), conflicting results have been reported, showing association of the rs29001323 TT genotype with either good clinical response and improved OS [44,45] or the opposite effect [43].

Another study has also explored the role of polymorphisms in two genes involved in the repair of cisplatin-induced interstrand crosslinks: DNA polymerase REV1 (DNA directed) (REV1) and REV3-like polymerase (DNA directed) zeta catalytic subunit (REV3L), in tumor DNA of 66 osteosarcoma patients [46]. Inferior EFS and OS were observed for patients harboring the REV1 variant rs3087403, as well as in a haplotype with the REV1L variant rs462779, suggesting a role for the REV genes in the efficacy of cisplatin treatment. Furthermore, a Nibrin (NBN) haplotype was also associated with EFS, although not with OS [26]. Although these studies are still in need of independent validation, these findings point towards a role of genetic variation in multiple DNA repair genes in the treatment efficacy primarily of cisplatin.

Pathway-based studies
A few pharmacogenetic studies have taken a broader approach than candidate gene analysis and have investigated multiple genes known to be involved in the metabolism and transport of cisplatin and doxorubicin-based chemotherapy. In addition to the study by Caronia et al. reporting on the role of ABC transporters discussed above, two other pathway-based studies have been published. Windsor et al. have analyzed 35 variants in 21 genes from the pathways of MAP chemotherapy in relation to HR and progression-free survival (PFS) in a cohort of 50 patients with osteosarcoma [23]. In addition to the results for GSTP1 discussed above, the variant ATP-binding cassette sub-family C member 2 (ABCC2) 24C>T (rs717620) was associated with a poor HR, whereas the variant 1958G>A (rs1950902) in methylenetetrahydrofolate dehydrogenase, cyclohydrolase and formyltetrahydrofolate synthetase 1 (MTHFD1) had a protective effect. However, the HR was not associated with PFS, making it difficult to interpret the clinical importance of the findings. No variants were found to be associated with PFS when corrected for the primary tumor site and metastatic status at diagnosis. However, the overrepresentation of survivors, because the study included only patients alive at the time of retrospective collection, might have limited the ability to detect associations with small effect sizes.

A complete pathway-based approach was taken in a previous study from our group [42]. All genes known to be involved in the metabolism and transport pathways of cisplatin and doxorubicin were investigated, including PD-related genes (e.g., DNA repair and apoptosis related genes). A total of 384 SNPs in 54 candidate genes were genotyped for two cohorts of a total of 177 patients with osteosarcoma. In a meta-analysis, five variants were identified that were significantly associated with 5-year PFS: Fas ligand (Fasl) rs763110, MutS homologue 2 (MSH2) rs4638843, caspase 3 (CASP3) rs2720376, ATP-binding cassette sub-family C member 5 (ABCC5) rs939338, and cytochrome P450 3A4 (CYP3A4) rs4646437. These genes were hypothesized to influence chemotherapy efficacy through differences in PD or in PK (the latter two). Risk stratification based on the combined effects of the risk alleles improved the ability to predict the 5-year PFS. Although this is the largest and most extensive study published to date with respect to patient numbers and coverage of metabolic pathways, validation in independent, larger cohorts is needed.

On the basis of all of the above studies into drug efficacy, it currently remains difficult to draw firm conclusions regarding the most important genes deserving further follow-up in prospective and functional validation studies, despite numerous candidate gene studies and several pathway-based studies. Large-scale studies with respect to numbers of patients and genes may be useful to pinpoint the role of these and previously not investigated genes in drug efficacy in patients with osteosarcoma.

Pharmacogenetics of drug toxicity
The ability to manage toxicities induced by chemotherapy is an important factor for the success of cancer treatment. Osteosarcoma patients receiving MAP therapy can experience severe side effects that are often dose-limiting. General side effects include cell-damaging effects, such as nausea and vomiting, myelosuppression, and immunosuppression. More specific side effects include nephrotoxicity, neurotoxicity, hepatotoxicity, hearing loss (ototoxicity) and cardiotoxicity. Pharmacogenetic markers can be a useful tool to predict the patient’s drug exposure and sensitivity to toxicities. With upfront knowledge on a patient’s susceptibility to specific side effects, alternative strategies or the use of protective agents can be considered to prevent treatment-related morbidity and mortality, and greatly improve the patient’s quality of life. Here we focus on the side effects of each of the chemotherapeutics in the MAP regimen that are studied in pharmacogenetic research in patients with osteosarcoma. An overview of the published studies is provided in Tables 2-4.

Cisplatin-induced ototoxicity
Ototoxicity is a major dose-limiting side effect of cisplatin. It occurs in 40-60% of cisplatin-treated patients and is characterized by permanent, bilateral, sensorineural hearing loss [47]. There is substantial interindividual variation in ototoxicity, which is not fully
explained by clinical risk factors such as young age at diagnosis, cumulative cisplatin dose and cranial irradiation [48-50]. This has led to the hypothesis that variation in genes encoding drug metabolizing enzymes and transporters could also be determinants of the sensitivity to cisplatin-induced hearing loss.

Six small-scale candidate gene pharmacogenetic studies including osteosarcoma patients have investigated polymorphisms in candidate genes mostly based on their expression or function in the inner ear. The presumed mechanism underlying cisplatin-induced hearing loss involves DNA damage and sensitivity of the cochlear cells to oxidative stress by the generation of toxic levels of reactive oxygen species (ROS) [51]. Accumulation of ROS leads to the depletion of the cochlear antioxidant enzyme system, including the GST enzymes. Peters et al. were the first to study genetic polymorphisms in five GST genes (GSTM1, GSTM3, GSTT1, GSP1, and GSTZ1) in a cohort of 39 cisplatin-treated young adults including 27 patients with osteosarcoma [52]. They have reported an increased frequency of the GSTM3*B allele in patients without hearing loss, indicating a protective effect. However, other studies could not detect associations of GST variants and ototoxicity, including a Brazilian study of 80 patients with osteosarcoma assessing GSTM3*B and variants of GSTM1 and GSTT1 [27,53].

Riedemann et al. have studied Megalin, which is expressed by the marginal cells of the cochlea, encoding a multiligand endocytic receptor that is implicated in binding cisplatin because of its association with uptake of aminoglycosides [54]. In 50 pediatric cancer patients (N=38 with osteosarcoma) the variant rs2075252 was found to be associated with ototoxicity after treatment. In contrast, a Canadian study could not identify a significant association of cisplatin-induced ototoxicity and the Megalin rs2075252 variant [53].

A fourth study has focused on NER genes in a cohort of 32 osteosarcoma patients, demonstrating an association of the rs2228001 SNP in the XPC gene with ototoxicity [33]. The authors suggested that the variant may decrease the DNA repair capacity, rendering outer hair cells in the inner ear susceptible to cisplatin.

Spracklen et al. have assessed the role of polymorphisms in the Otos gene [55]. This gene encodes otospirallin, which is a protein involved in cochlear function and protection and has been implicated in the response to cisplatin in cell models [56]. In a cohort of 100 cisplatin-treated adult cancer patients of varying ethnicity (N=10 with osteosarcoma), two SNPs (rs77124181 and rs2291767, in full LD in Caucasians) were found to be protective. Validation of this finding is needed, especially as the cohort was heterogeneous with respect to dosing and schedule, and as all osteosarcoma patients were reported to have developed ototoxicity.

Another recent candidate gene study has been reported by Lanvers-Kaminsky et al., investigating the solute carrier family 31 member 1 (SLC31A1) gene and solute carrier family 22 member 2 (SLC22A2), the latter because of its involvement in mediating the transport of cisplatin [57]. They determined that SLC22A2 variant rs316019 was
Cardiotoxicity is a major problem in the treatment with the anthracycline doxorubicin. It can present as acute, chronic, or late onset heart failure and appears to be induced by oxidative stress in cardiomyocytes [64]. There is large variation in the susceptibility to anthracycline-induced cardiotoxicity (ACT) and the risk for ACT is increased with younger age, higher cumulative dose, and female sex [65]. However, these factors are not sufficient to accurately predict a patient’s sensitivity to developing ACT. Therefore, pharmacogenetic studies have interrogated 1,949 SNPs in 220 key genes in a discovery cohort and a replication cohort with a total of 162 pediatric cancer patients, including 40 patients with osteosarcoma [65]. There are large numbers of genes and still have sufficient power in a relatively small patient cohort. Larger effect sizes are expected. This makes it possible to investigate drug efficacy reported so far, which could be inherent to the outcome, because generally in toxicity studies larger effect sizes are expected. This also makes it possible to investigate in an independent cohort of 153 patients [68]. However, these findings could not be consistently replicated to large heterogeneity of patient cohorts with respect to ethnicity and different cancers, the latter reflected by variation in treatment regimens. The controversy has been the subject of a debate, especially because the original findings led to a cisplatin label change by the US Food and Drug Administration, indicating the association of cisplatin-induced hearing loss in a discovery cohort of pediatric patients and replication cohort with a total of 41 osteosarcoma patients, and an adult cancer replication cohort.

### General introduction

1. **Consistency replicated by ourselves and others**

2. **Ref.**

<table>
<thead>
<tr>
<th>No. of osteosarcoma patients (of total)</th>
<th>Ethnic origin nationality</th>
<th>Drugs**</th>
<th>Genes investigated</th>
<th>Associated variant(s)/?</th>
<th>Possible functional role of associated variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=80 White, Asian, Brown/unspecified; Brazilian</td>
<td>CAF1</td>
<td>GSTM1, GSTM3, GSTT1</td>
<td>None</td>
<td>NA</td>
<td><strong>[27]</strong></td>
</tr>
<tr>
<td>N=50 Caucasian, Afro-Caribbean, Indian / Asian, English</td>
<td>MAP 21 genes from MAP pharmacological pathways</td>
<td>GSTP1 rs1695</td>
<td>Enzyme activity ↓</td>
<td><strong>[23]</strong></td>
<td></td>
</tr>
<tr>
<td>N=79 (of 407) Non-Hispanic white, Hispanic, Black, other unspecified; American</td>
<td>A, other NS</td>
<td>CB90, CB93</td>
<td>CB93 rs1056892</td>
<td><strong>[64]</strong></td>
<td></td>
</tr>
<tr>
<td>N=27 (of 448) NS, Canadian, Dutch</td>
<td>A, other NS</td>
<td>220 drug metabolism genes</td>
<td>SLC28A3 rs7853758</td>
<td><strong>[67]</strong></td>
<td></td>
</tr>
<tr>
<td>N=16 (of 202) NS, Canadian, Dutch</td>
<td>A, other NS</td>
<td>15 genes: ABC1, ABC2, ABC4, ABC7, ABCC1, CYP2F1, FMO2, FMO3, HNMT, SLC10A2, SLC22A2, SLC28A1, SLC28A3</td>
<td>Unknown</td>
<td><strong>[68]</strong></td>
<td></td>
</tr>
<tr>
<td>N=54 (of 363) Non-Hispanic white, other unspecified; American</td>
<td>A, other NS</td>
<td>2100 cardiovascular genes</td>
<td>MAS3 rs1253228</td>
<td><strong>[69]</strong></td>
<td></td>
</tr>
<tr>
<td>N=35 (of 530) NS, Canadian, Dutch</td>
<td>A, other NS</td>
<td>&gt;1000 key drug interaction genes</td>
<td>SLC22A7 rs4144978 SLC22A7 rs4882753</td>
<td>Unknown</td>
<td><strong>[70]</strong></td>
</tr>
<tr>
<td>N=20 (of 376) NS, Canadian, Dutch</td>
<td>A, other NS</td>
<td>Germline-wide</td>
<td>CMG6 rs2229774</td>
<td>RARG activity ↓</td>
<td><strong>[71]</strong></td>
</tr>
</tbody>
</table>

---

**Notes:**
- **Table 3.** Characteristics of pharmacogenetic studies of cardiotoxicity in patients with osteosarcoma.
- **Drugs column** indicates drugs included in treatment of the osteosarcoma patients as far as stated; drugs between brackets are not given to all patients in the study.
- **Ref.**

---

**General introduction:**

- Number of bone tumors; osteosarcoma not specified.
- For the large scale pathway and genome-wide studies, only the associated variants significant after correction for multiple testing and/or replication within the same study are included in the table.
In recent report from Visscher et al., the authors aimed to identify additional genetic variants that might have been missed in their previous study, by investigating over 300 genes in a discovery cohort and a replication cohort.

In a two-stage design, they included a total of 2,100 genes investigated in relation to ACT. However, the replication cohort included in the original publication did not report association of variants in solute carrier family 28 member 3 (SLC28A3) and UDP glucuronosyltransferase family 1 member A6 (UGT1A6) genes with ACT, although the effect was only significant when the two Canadian cohorts of the original study and replication cohorts were combined, and the replication cohort also encompassed, in part, patients of the Dutch replication cohort included in the original publication.

A large study of 487 pediatric cancer patients has investigated polymorphisms in cytochrome P450 (CYP) and solute carrier family 22 (SLC) genes in a discovery cohort of 198 and a Dutch replication cohort of 289 pediatric cancer patients in combination with a number of other genes with ACT. An additional replication study from the same group has reported confirmation of variants in SLC22A7 and SLC28A3 and a number of other genes with ACT in a Canadian discovery cohort of 156 pediatric cancer patients, a Canadian replication cohort (N=188) and a Dutch replication cohort (N=96). Visscher et al. did not report on associations of variants in solute carrier family 28 member 3 (SLC28A3) and UDP glucuronosyltransferase family 1 member A6 (UGT1A6) in association with ACT, although not when patients were treated with high doses as used in the treatment of osteosarcoma (cumulative dose 450 mg/m^2). However, the CYP3A4*1B and CYP3A5*3A variants were not found to be associated in another large replication study published simultaneously, in which 762 drug metabolism and transport genes were interrogated in a Canadian discovery cohort of 218 pediatric cancer patients and 283 healthy controls.

Table 4. Characteristics of pharmacogenetic studies of MTX-induced toxicities in patients with osteosarcoma.

<table>
<thead>
<tr>
<th>No. of osteosarcoma patients (of total)</th>
<th>Ethnic origin; nationality</th>
<th>Drugs*</th>
<th>Type of MTX-induced toxicity</th>
<th>Genes investigated</th>
<th>Associated variant(s)</th>
<th>(Possible) functional role of associated variant</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=50</td>
<td>Caucasian, Mio-Caribian,</td>
<td>MAP</td>
<td>Hematologic, mucositis</td>
<td>21 genes from MAP pharmacological pathways</td>
<td>MTHFR rs1801131, A BCB1 1044592, ABC2 rs1722732</td>
<td>MTHFR activity ↓, ABCB1 efflux activity ↓, ABC2 efflux activity ↓</td>
<td>[23]</td>
</tr>
<tr>
<td>N=62</td>
<td>NS, Norwegian</td>
<td>MAP(2)</td>
<td>Hepatotoxicity, mucositis</td>
<td>ABCC2, ZMTR, MTHFR, RFC1</td>
<td>MTHFR rs1801135, Z MTR rs1650723</td>
<td>MTHFR activity ↓, ZMTR activity ↓</td>
<td>[25]</td>
</tr>
<tr>
<td>N=40</td>
<td>NS, Dutch</td>
<td>MAP(2)</td>
<td>Hepatotoxicity</td>
<td></td>
<td></td>
<td>MTHFR activity ↓</td>
<td>[25]</td>
</tr>
<tr>
<td>N=41†</td>
<td>Caucasian, Slovenian</td>
<td>MAP</td>
<td>Hematologic, hepato-, nephro- and GI toxicity</td>
<td>ABCB1, ABC2, MTHFR, NTR, MTR, MTX, SLC29A1, SLC12A1</td>
<td>MTHFR rs717620</td>
<td>MTHFR activity ↓, not in meta-analysis</td>
<td>[72]</td>
</tr>
<tr>
<td>N=96</td>
<td>NS, Spanish</td>
<td>MAP(2),</td>
<td>Hematologic, hepato-, nephro- and GI toxicity</td>
<td>ABCB1, ABC2, MTHFR, NTR, MTX, MTR, SLC29A1, SLC12A1</td>
<td>MTHFR rs1801135</td>
<td>MTHFR activity ↓, not in meta-analysis</td>
<td>[74]</td>
</tr>
<tr>
<td>N=7</td>
<td>NS, Turkish</td>
<td>MTX, CYT, CHS</td>
<td>Biomedical or clinical toxicities</td>
<td>MTRR, SLC29A1, SLC12A1</td>
<td>MTHFR rs1801135</td>
<td>MTHFR activity ↓, not in meta-analysis</td>
<td>[73]</td>
</tr>
<tr>
<td>N=37</td>
<td>Asian, Roman</td>
<td>MAP(2)</td>
<td>Hepato-, nephro- and neurotoxicity, mucositis</td>
<td>ATC, MTHFR, RFC1</td>
<td>RFC1 rs1581266</td>
<td>MTHFR activity ↓</td>
<td>[76]</td>
</tr>
</tbody>
</table>

NS, not specified; GI, gastrointestinal; M, methotrexate; A, adriamycin (doxorubicin); P, cisplatin; I, ifosfamide; E, etoposide; VCR, vincristine; CTX, cyclophosphamide; ACD, actinomycin D.

* Drugs column indicates drugs included in treatment of the osteosarcoma patients as far as stated; drugs between brackets are not given to all patients in the study.

† No germline, but tumor DNA was analyzed.

§ Studies into the functional effect of the genetic variant have reported conflicting results.
in case of sustained high plasma levels, to minimize toxicities. However, leucovorin rescue may potentially reduce the desired antitumor effect depending on the dosage and the time point it is given. Upright knowledge of the patient’s response to MTX could therefore be helpful in the prevention of MTX-induced toxicities and their consequences.

Several pharmacogenetic studies have suggested a role of genetic variants in the interpatient variability of the response to high-dose (HD)-MTX, of which a few have included osteosarcoma patients. Windsor et al., taking a MAP pathway-based approach in 50 patients with osteosarcoma by selecting genes of the metabolic pathways of methotrexate, doxorubicin, and cisplatin, have suggested associations of variants of methylene tetrahydrofolate reductase (MTHFR) (1298A>C, rs1801131), ABCB1 (3435T>C, rs1045642), and ABCC2 (3563T>A, rs1722723) with anemia, leucopenia or mucositis as a result of HD-MTX treatment [23]. ABCB1 and ABCC2 are known transporters of MTX; however, the role of polymorphisms in the genes in the context of MTX treatment is not well established in osteosarcoma and other cancers, although another variant of ABCC2 was also related to MTX-induced toxicities in another small scale study of tumor DNA of 44 patients with osteosarcoma [72]. MTHFR encodes a key enzyme in the folate homeostasis and metabolism and is more widely studied in MTX-treated patients in relation to hepatotoxicity. Dogan et al. have reported an increased frequency of MTHFR variant 667C>T (rs1801133) in patients with osteosarcoma experiencing high alanine aminotransferase when compared to the general population, although no statistics could be performed owing to the very small number of patients (N=7) [73]. Jabeen et al. have recently reported an association of the same variant with hepatotoxicity in a candidate gene study of 62 patients with osteosarcoma [25]. However, the association of the 667C>T variant and hepatotoxicity could not be confirmed by other studies, including a large meta-analysis also including patients with osteosarcoma [72,74-76]. Hence, the role of variants in this gene in MTX-induced toxicity still needs to be elucidated [77]. The RFC1 rs1051266 variant that was identified by Jabeen et al. in relation to survival after MTX-based therapy was not associated with MTX-induced toxicities in the same study [25]; however, a Korean study showed association of the variant with MTX serum levels and mucositis, indicating that this variant deserves follow-up studies to further elucidate its role not only with respect to survival but also toxicity after MTX treatment [76].

In the candidate gene study of 44 osteosarcoma patients by Goricar et al., associations of a solute carrier organic anion transporter family member 1B1 (SLCO1B1) variant were found to influence MTX serum levels, although this did not translate into associations with MTX-induced toxicities [72]. Other variants in this organic anion transporter gene were previously found in association with MTX clearance in a large GWAS in patients with acute lymphoblastic leukemia, which was also validated by others [78,79]. Further validation in large cohorts of osteosarcoma patients is warranted to confirm the influence of SLCO1B1 variants on MTX PK and susceptibility to toxicities in this group of patients.

The next step

Most of the pharmacogenetic studies in osteosarcoma performed to date, especially with respect to drug efficacy, have used a candidate gene approach by design, investigating a limited number of genes involved in pharmacokinetic or pharmacodynamic pathways of MAP therapy. These studies have yielded promising results, strongly indicating a role of multiple genetic polymorphisms in the efficacy of chemotherapeutic treatment or sensitivity to toxicities. A logical next step is to take a broader approach, as certainly not all variation in the treatment efficacy and toxicity can be explained by the genetic variants studied and identified so far. In addition, the genetic component of the treatment response is likely to be a combination of multiple small genetic effects.

Aim and outline of this thesis

The overall aim of this thesis is to identify genetic predictors of efficacy and toxicity of chemotherapeutic treatment in patients with osteosarcoma, which could ultimately be used to establish a prediction model to enable risk stratification and tailored treatment through optimization, or which could serve as new therapeutic targets for osteosarcoma treatment.

All of the studies presented in this thesis involved patients with high-grade osteosarcoma. In the majority of the studies included in this thesis, we have used the comprehensive Drug Metabolizing Enzymes and Transporters (DMET) Plus array to assess genetic variants in genes related to drug absorption, distribution, metabolism and transport, in retrospectively collected osteosarcoma patients cohorts. As for a proportion of patients the only available source for germline DNA was normal formalin-fixed, paraffin-embedded (FFPE) tissue, the performance and reliability of FFPE tissue-derived DNA on the DMET Plus array was examined, which is described in chapter 2.

The next two chapters are focused on the pharmacogenetics of chemotherapy-induced toxicities. In chapter 3, we report on the association of a variant in the ACYP2 gene with cisplatin-induced ototoxicity, which was previously identified in a GWAS in children with brain tumors, and confirmed in our osteosarcoma patient population. In chapter 4, we describe a study into the role of genetic variants in MTX-induced hepatotoxicity using the DMET Plus array.

Chapters 5-7 are focused on the pharmacogenetics of chemotherapy efficacy. In a previous pathway-based study by our group, five genetic markers were identified that could distinguish patients with good outcome from those with poor outcome. When combined in a genetic risk score, these markers were predictive of 5-year PFS. In chapter 5, we report on the validation of this genetic risk score in independent osteosarcoma patient cohorts. In chapter 6, we have employed the DMET Plus array to detect (additional) genetic
variants associated with PFS and HR in our Dutch patient cohorts and collaborative cohorts. In chapter 7, we report on a DMET analysis focusing on the subgroup of patients with an inadequate drug response (progressive disease) already during first-line chemotherapeutic treatment.

In the final chapter, the research described in this thesis is summarized and future research perspectives are given.

References


Methods
Chapter 2

High-quality genotyping data from formalin-fixed, paraffin-embedded tissue on the drug metabolizing enzymes and transporters plus array

Hanneke I. Vos, Tahar van der Straaten*, Marieke J.H. Coenen*, Uta Flucke, D. Maroeska W.M. te Loo, Henk-Jan Guchelaar

* These authors contributed equally

Journal of Molecular Diagnostics 2015 Jan;17(1):4-9
Abstract

The Affymetrix Drug Metabolizing Enzymes and Transporters (DMET) Plus array covers 1,936 markers in 231 genes involved in drug metabolism and transport. Blood- and saliva-derived DNA works well on the DMET array, but the utility of DNA from formalin-fixed, paraffin-embedded (FFPE) tissue has not been reported for this array. As the ability to use DNA from FFPE tissue on the array could open the potential for large retrospective sample collections, we examined the performance and reliability of FFPE-derived DNA on the DMET Plus array.

Germline DNA isolated from archived normal FFPE tissue blocks stored for 3 to 19 years and matched blood or saliva from 16 patients with osteosarcoma was genotyped on the DMET Plus array. Concordance was assessed by calculating agreement and the κ-statistic.

We observed high call rates for both the blood- or saliva-derived DNA samples (99.4%) and the FFPE-derived DNA samples (98.9%). Moreover, the concordance among the sixteen blood- or saliva-derived DNA and FFPE DNA pairs was high (97.4%, κ=0.915).

This is the first study showing that DNA from normal FFPE tissue provides accurate and reliable genotypes on the DMET Plus array compared with blood- or saliva-derived DNA. This finding provides an opportunity for pharmacogenetic studies in diseases with high mortality rates and prevents a bias in studies where otherwise only alive patients can be included.

Introduction

The contribution of genetic variation to interindividual variability of drug efficacy and toxicity profiles is widely acknowledged. The Affymetrix Drug Metabolizing Enzymes and Transporters (DMET) Plus array is a dedicated assay for pharmacogenetic applications as it genotypes 1,936 markers in 231 genes involved in drug absorption, distribution, metabolism, excretion and transport [1,2]. The array is based on molecular inversion probe (MIP) technology for the simultaneous genotyping of multiple polymorphisms in a single assay. Blood and saliva samples are used as the standard source of germline DNA for genotyping [3]. However, collection and storage of germline DNA have not routinely been performed in most clinical trials in the past. In particular, in pharmacogenetic studies in the field of oncology with cancers with low survival rates, the lack of stored blood or saliva can result in biased patient populations for research because the poor responders to therapy are likely to be missing. In such cases, stored formalin-fixed, paraffin-embedded (FFPE) tissue is often the only available source for germline DNA. However, in the process of fixation and storage of FFPE samples, DNA degradation and DNA-protein cross-linkages result in reduced DNA quality, which can be challenging for PCR-based analysis [4,5]. Therefore, the use of archived tissue, which can be highly valuable for biomarker studies, first needs to be validated for the specific analysis [6].

For several Affymetrix single nucleotide polymorphism (SNP) arrays, the performance of FFPE-derived DNA has been compared with DNA from fresh frozen tissue, revealing a high degree of concordance; however, these studies have primarily focused on copy number variations and chromosomal aberrations in tumor tissue [7-9]. In addition, Affymetrix has successfully demonstrated the suitability of FFPE-derived DNA for MIP technology using both normal and tumor FFPE samples [10]. DNA from FFPE tissue may also be suitable for genotyping by the DMET array; however, the DMET Plus array differs from standard MIP arrays because it performs an initial 36-plex multiplex PCR to preamplify genes known to have pseudogenes or close homologs. To date, no studies have reported the utility of DNA from FFPE samples as a template for the DMET Plus array, and no comparison of genotype performance with DNA from blood or saliva has been made. In the present study, the performance and reliability of FFPE-derived DNA on the DMET Plus array are assessed using a series of matched blood or saliva and normal FFPE samples from patients with osteosarcoma.
Materials and methods

DNA extraction
From 16 patients with primary, high-grade osteosarcoma treated at the Radboud university medical center (Nijmegen, The Netherlands), DNA was isolated from archived normal FFPE tissue and blood or saliva after obtaining written informed consent from the patients or parents. If clinically available, blood was used as the primary source for reference DNA; otherwise, saliva was used. The study was approved by the local ethics committee, and all tissue was handled according to the Code Proper Use of Human Tissue in The Netherlands (FEDERA, Federation of Dutch Medical Scientific Societies; www.federa.org; last accessed April 16, 2014).

DNA was isolated from fresh EDTA whole blood samples (n=13) using the QIAamp DNA Blood Midi kit (Qiagen, Venlo, The Netherlands) according to the manufacturer’s protocol. Saliva was collected (n=3) using the Oragene DNA saliva collection kit (DNA Genotek Inc., Kanata, Ontario, Canada), and DNA was extracted following the manufacturer’s instructions.

FFPE tissue-derived DNA from the same patients was obtained using stored FFPE tissue blocks from surgical resection specimens. At the time of resection, fresh tissue was fixed in 4% buffered formalin, processed routinely, and embedded in paraffin as described previously [11]. At the time of DNA isolation, the FFPE tissue blocks had been stored at room temperature for 3 to 19 years. Before DNA isolation, all tissue blocks were assessed by an experienced pathologist (U.F.) using hematoxylin and eosin-stained slides. From a series of tissue blocks with optically distinguishable normal and/or malignant tissue, FFPE tissue blocks that contained only normal tissue were selected to ensure isolation of germline DNA only. All tissue blocks contained nondecalcified, normal soft tissue, including lung, muscle, skin, lymph node, or a combination. The FFPE tissue samples were sectioned using a microtome, which was cleaned with ethanol between cases to avoid contamination. Ten 20 µm tissue sections were used for isolation as samples were sectioned using a microtome, which was cleaned with ethanol between samples to avoid contamination. Ten 20 µm tissue sections were used for isolation as samples were sectioned using a microtome, which was cleaned with ethanol between

Quality control of DNA samples
To assess the quality of the DNA samples, a qualitative size range PCR was performed as described previously [11]. In brief, 50 ng of genomic DNA (based on NanoDrop measurements using Nanodrop 2000; Thermo Scientific, Breda, The Netherlands) from FFPE tissue (n=16) and one representative blood sample was included. The assay contained five primers sets that amplify fragments that range from 100 to 400 bp, as described previously [12]. MilliQ was included as negative control and 50 ng DNA from frozen tissue sections as a positive control for the assay. The PCR products were visualized and recorded in an ethidium bromide-stained 2% agarose gel using the Imago (Isogen Bioscience, IJsselstein, The Netherlands).

Genotyping
Thirty-two DNA samples were genotyped using the DMET Plus array according to the manufacturer’s instructions (Affymetrix UK Ltd, High Wycombe, United Kingdom) [1]. All samples were normalized to 60 ng/µL based on Nanodrop measurements, and 17 µl (1.02 µg DNA) per sample was used for genotyping. Genotypes were calculated with DMET console software 1.3 using Dynamic Genotype Boundaries algorithm version 2.

In addition, the 16 blood- or saliva-derived DNA samples were genotyped for 381 markers, of which 13 SNPs were overlapping with the DMET Plus panel, using a custom-based assay on the Illumina BeadArray platform (Illumina Inc., San Diego, CA). Genotypes were calculated using Illumina Genome Studio software version 2008.1 (Illumina Inc.).

Statistical analysis
Genotyping results of the 16 DNA pairs from blood or saliva and FFPE tissue were evaluated for call rate and SNP and copy number variation concordance. Call rates were derived from the DMET console software calculation, which is the percentage of SNPs successfully genotyped per sample, excluding no calls but including possible rare allele (PRA) calls. PRA calls are assigned by the DMET console software to probe signals that fall out of the range of the training data because the variant alleles were rarely or not observed in assay development. For each marker, the concordance between the genotypes obtained from the blood- or saliva-derived DNA and the FFPE-derived DNA was determined by calculation of the percentage of genotype agreement and the κ-statistic. The κ-statistic measures agreement beyond chance alone; a κ-statistic of 0 indicates chance agreement, whereas a κ-statistic of 1 indicates complete agreement. The overall agreement and overall κ-statistic were calculated as the mean of the agreement and κ of all markers. The analyses were performed using R version 3.0.2 (http://www.r-project.org) and included no calls and PRA calls. In addition, genotype concordance was calculated for each DNA pair with and without inclusion of PRA calls (i.e. the percentage of markers with the same genotype call for DNA from blood or saliva and FFPE tissue

Genotyping FFPE tissue DNA on DMET array

Statistical analysis
Genotyping results of the 16 DNA pairs from blood or saliva and FFPE tissue were evaluated for call rate and SNP and copy number variation concordance. Call rates were derived from the DMET console software calculation, which is the percentage of SNPs successfully genotyped per sample, excluding no calls but including possible rare allele (PRA) calls. PRA calls are assigned by the DMET console software to probe signals that fall out of the range of the training data because the variant alleles were rarely or not observed in assay development. For each marker, the concordance between the genotypes obtained from the blood- or saliva-derived DNA and the FFPE-derived DNA was determined by calculation of the percentage of genotype agreement and the κ-statistic. The κ-statistic measures agreement beyond chance alone; a κ-statistic of 0 indicates chance agreement, whereas a κ-statistic of 1 indicates complete agreement. The overall agreement and overall κ-statistic were calculated as the mean of the agreement and κ of all markers. The analyses were performed using R version 3.0.2 (http://www.r-project.org) and included no calls and PRA calls. In addition, genotype concordance was calculated for each DNA pair with and without inclusion of PRA calls (i.e. the percentage of markers with the same genotype call for DNA from blood or saliva and FFPE tissue

Genotyping FFPE tissue DNA on DMET array

Statistical analysis
Genotyping results of the 16 DNA pairs from blood or saliva and FFPE tissue were evaluated for call rate and SNP and copy number variation concordance. Call rates were derived from the DMET console software calculation, which is the percentage of SNPs successfully genotyped per sample, excluding no calls but including possible rare allele (PRA) calls. PRA calls are assigned by the DMET console software to probe signals that fall out of the range of the training data because the variant alleles were rarely or not observed in assay development. For each marker, the concordance between the genotypes obtained from the blood- or saliva-derived DNA and the FFPE-derived DNA was determined by calculation of the percentage of genotype agreement and the κ-statistic. The κ-statistic measures agreement beyond chance alone; a κ-statistic of 0 indicates chance agreement, whereas a κ-statistic of 1 indicates complete agreement. The overall agreement and overall κ-statistic were calculated as the mean of the agreement and κ of all markers. The analyses were performed using R version 3.0.2 (http://www.r-project.org) and included no calls and PRA calls. In addition, genotype concordance was calculated for each DNA pair with and without inclusion of PRA calls (i.e. the percentage of markers with the same genotype call for DNA from blood or saliva and FFPE tissue

Genotyping FFPE tissue DNA on DMET array

Statistical analysis
Genotyping results of the 16 DNA pairs from blood or saliva and FFPE tissue were evaluated for call rate and SNP and copy number variation concordance. Call rates were derived from the DMET console software calculation, which is the percentage of SNPs successfully genotyped per sample, excluding no calls but including possible rare allele (PRA) calls. PRA calls are assigned by the DMET console software to probe signals that fall out of the range of the training data because the variant alleles were rarely or not observed in assay development. For each marker, the concordance between the genotypes obtained from the blood- or saliva-derived DNA and the FFPE-derived DNA was determined by calculation of the percentage of genotype agreement and the κ-statistic. The κ-statistic measures agreement beyond chance alone; a κ-statistic of 0 indicates chance agreement, whereas a κ-statistic of 1 indicates complete agreement. The overall agreement and overall κ-statistic were calculated as the mean of the agreement and κ of all markers. The analyses were performed using R version 3.0.2 (http://www.r-project.org) and included no calls and PRA calls. In addition, genotype concordance was calculated for each DNA pair with and without inclusion of PRA calls (i.e. the percentage of markers with the same genotype call for DNA from blood or saliva and FFPE tissue

Statistical analysis
Genotyping results of the 16 DNA pairs from blood or saliva and FFPE tissue were evaluated for call rate and SNP and copy number variation concordance. Call rates were derived from the DMET console software calculation, which is the percentage of SNPs successfully genotyped per sample, excluding no calls but including possible rare allele (PRA) calls. PRA calls are assigned by the DMET console software to probe signals that fall out of the range of the training data because the variant alleles were rarely or not observed in assay development. For each marker, the concordance between the genotypes obtained from the blood- or saliva-derived DNA and the FFPE-derived DNA was determined by calculation of the percentage of genotype agreement and the κ-statistic. The κ-statistic measures agreement beyond chance alone; a κ-statistic of 0 indicates chance agreement, whereas a κ-statistic of 1 indicates complete agreement. The overall agreement and overall κ-statistic were calculated as the mean of the agreement and κ of all markers. The analyses were performed using R version 3.0.2 (http://www.r-project.org) and included no calls and PRA calls. In addition, genotype concordance was calculated for each DNA pair with and without inclusion of PRA calls (i.e. the percentage of markers with the same genotype call for DNA from blood or saliva and FFPE tissue

Statistical analysis
Genotyping results of the 16 DNA pairs from blood or saliva and FFPE tissue were evaluated for call rate and SNP and copy number variation concordance. Call rates were derived from the DMET console software calculation, which is the percentage of SNPs successfully genotyped per sample, excluding no calls but including possible rare allele (PRA) calls. PRA calls are assigned by the DMET console software to probe signals that fall out of the range of the training data because the variant alleles were rarely or not observed in assay development. For each marker, the concordance between the genotypes obtained from the blood- or saliva-derived DNA and the FFPE-derived DNA was determined by calculation of the percentage of genotype agreement and the κ-statistic. The κ-statistic measures agreement beyond chance alone; a κ-statistic of 0 indicates chance agreement, whereas a κ-statistic of 1 indicates complete agreement. The overall agreement and overall κ-statistic were calculated as the mean of the agreement and κ of all markers. The analyses were performed using R version 3.0.2 (http://www.r-project.org) and included no calls and PRA calls. In addition, genotype concordance was calculated for each DNA pair with and without inclusion of PRA calls (i.e. the percentage of markers with the same genotype call for DNA from blood or saliva and FFPE tissue
of the total number of markers, excluding those with no calls and/or PRA for DNA from either source). Calculations per pair were performed using Microsoft Excel (Microsoft Corp, Redmond, WA).

Assessment of associations of SNPs with DNA source (blood or saliva versus FFPE tissue) was performed by logistic regression analysis using PLINK software version v1.07 (http://pngu.mgh.harvard.edu/purcell/plink; last accessed April 14, 2014) [13]. P < 0.05 was used as significance level, without correction for multiple testing because of the exploratory nature of this analysis.

Results

To assess the suitability of DNA isolated from FFPE tissue for the DMET Plus array, germline DNA derived from FFPE tissue and matched blood or saliva samples as high quality reference material was collected from 16 patients with osteosarcoma. The average yield of the FFPE-derived DNA samples was 12.2 µg (range, 1.93 to 36.8 µg); hence, isolation yielded sufficient DNA from all FFPE samples for downstream array analysis.

Quality of DNA isolated from FFPE tissue

Prior to DMET analysis, the quality of all FFPE tissue-derived DNA samples was evaluated by size range PCR. The reference blood DNA revealed bands up to 400 bp using gel electrophoresis, whereas the FFPE DNA samples showed amplicon sizes ranging from 100 to 300 bp, with a reduced intensity of the largest fragment visible in most cases (Figure 1).

Genotype call rates

All DNA samples from blood, saliva, and FFPE tissue were successfully genotyped by the DMET Plus array. The mean ± SD call rate for all blood- or saliva-derived DNA samples was 99.4% ± 0.30% (Table 1). For the FFPE-derived DNA samples, the mean ± SD call rate was 98.9% ± 1.0%. Fourteen FFPE DNA samples had call rates >96%, which is the threshold recommended by the manufacturer (Table 1) [1]. The two FFPE DNA samples with lower call rates (96.32% and 96.53%) were, however, still within a range suitable for genetic association studies, in which a minimum of 95% is considered acceptable. Although these samples were also observed to have small amplicon sizes in the size range PCR (maximum, 100 bp), indicative of reduced DNA quality, this correlation was not consistent among the other samples with similar fragment sizes. However, a reduction in the signal intensity of the 100-bp band was observed specifically in the two samples with lower call rates.

Table 1. Genotyping call rate and concordance per blood- or saliva-derived DNA and FFPE DNA pair.

<table>
<thead>
<tr>
<th>Sample no.</th>
<th>Blood or saliva Call rate, %</th>
<th>FFPE tissue Call rate, %</th>
<th>Excluding no calls Genotype concordance, %</th>
<th>Excluding no calls and PRAs Genotype concordance, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>99.68</td>
<td>99.38</td>
<td>99.48</td>
<td>99.84</td>
</tr>
<tr>
<td>2</td>
<td>99.74</td>
<td>99.07</td>
<td>99.37</td>
<td>99.89</td>
</tr>
<tr>
<td>3</td>
<td>99.59</td>
<td>99.64</td>
<td>99.48</td>
<td>99.84</td>
</tr>
<tr>
<td>4</td>
<td>99.07</td>
<td>99.69</td>
<td>99.37</td>
<td>99.89</td>
</tr>
<tr>
<td>5</td>
<td>99.22</td>
<td>99.43</td>
<td>99.42</td>
<td>99.89</td>
</tr>
<tr>
<td>6</td>
<td>99.33</td>
<td>96.32</td>
<td>97.46</td>
<td>98.42</td>
</tr>
<tr>
<td>7</td>
<td>98.81</td>
<td>99.27</td>
<td>99.26</td>
<td>99.79</td>
</tr>
<tr>
<td>8</td>
<td>98.91</td>
<td>98.55</td>
<td>98.78</td>
<td>99.25</td>
</tr>
<tr>
<td>10</td>
<td>99.48</td>
<td>99.38</td>
<td>99.16</td>
<td>99.68</td>
</tr>
<tr>
<td>11</td>
<td>99.48</td>
<td>99.43</td>
<td>98.90</td>
<td>99.53</td>
</tr>
<tr>
<td>12</td>
<td>99.53</td>
<td>99.22</td>
<td>98.85</td>
<td>99.31</td>
</tr>
<tr>
<td>13</td>
<td>99.53</td>
<td>98.76</td>
<td>98.84</td>
<td>99.26</td>
</tr>
<tr>
<td>14</td>
<td>99.90</td>
<td>99.17</td>
<td>98.90</td>
<td>99.37</td>
</tr>
<tr>
<td>15</td>
<td>99.12</td>
<td>96.53</td>
<td>96.54</td>
<td>98.97</td>
</tr>
<tr>
<td>16</td>
<td>99.69</td>
<td>99.74</td>
<td>99.48</td>
<td>99.69</td>
</tr>
</tbody>
</table>

Sample call rates and genotype concordance for DNA derived from blood (samples 1 to 13), saliva (samples 14 to 16), and matching FFPE tissue on the Affymetrix Drug Metabolizing Enzymes and Transporters Plus array.

FFPE, formalin-fixed, paraffin-embedded; PRA, possible rare allele.

Figure 1. Quality of DNA isolated from formalin-fixed, paraffin-embedded (FFPE) tissue. The size ranges of the PCR products of 16 DNA samples from FFPE tissue are given. Lane B is the representative DNA sample from blood (matching with FFPE sample no. 7); lane M, 100-bp ladder.
Genotype concordance

To analyze the reliability and accuracy of the genotype calling of FFPE-derived DNA samples by the DMET Plus array, the genotype concordance was calculated for the paired blood- or saliva- derived DNA and FFPE DNA genotyping results (Table 2). Of the 1931 markers, 1061 showed no genetic variation in our population, which is expected based on the known minor allele frequencies and previous DMET studies [14-16]. The genotype concordance was calculated with all markers and with only the 870 polymorphic markers. The overall agreement with all markers included was 97.4% with a κ-statistic of 0.915. The κ-statistic calculated per marker ranged from -0.143 (TPMT rs1800460, CYP21A2 rs72552757) to 1 (e.g. CHST3 rs4148943, FMO3 rs1736557), excluding the markers that were monomorphic, for which the κ-statistic was also defined as 1. Analysis of the concordance of the 35 genes with a very important pharmacogene (VIP) status (PharmGKB, http://www.pharmgkb.org, last accessed July 14, 2014) covered by the DMET Plus array revealed an overall concordance of 96.6% with a κ-statistic of 0.884, ranging from a mean of 68.8% for the markers in SLC19A1 to 100% in e.g. COMT and GSTT1 (Supplementary Table S1).

In addition, the genotype concordance was calculated per blood- or saliva-derived DNA and FFPE DNA pair to assess the potential link between concordance and FFPE DNA call rate (Table 1). Genotype concordance excluding no calls and PRAs exceeded 99.2% with the exception of the two pairs of which the FFPE DNA samples showed call rates <98%. Of seven markers miscalled in more than five blood- or saliva-derived DNA and FFPE DNA pairs, the cluster plots of four markers were considered unreliable. Without these markers, the genotype concordance would be somewhat higher; therefore, the reported concordances are slightly underestimated. Furthermore, the copy number calling was 100% concordant across the samples.

The genotypes obtained from the blood- or saliva-derived DNA samples were validated for a subset of 13 SNPs, which were included on the Illumina BeadArray platform as well. Genotypes of these SNPs (rs1045642, rs1128503, rs2231142, rs1801243, rs1005695, rs3787728, rs1056892, rs776746, rs1138272, rs1800566, rs89453, rs10898, rs4407290) were 100% concordant with the genotypes obtained using the DMET Plus array.

Discordant results

Of the 799 mismatches among the 30,896 SNP-sample pair comparisons, most (81.1%) involved no calls or PRA calls (i.e. switches from a genotype call in blood or saliva DNA to no call or PRA call in FFPE tissue DNA or vice versa and PRA call to no call or vice versa). The remaining genotype to genotype switches were dominated by switches of homozygous calls to heterozygous calls (11.01%), followed by heterozygous to homozygous switches (6.26%), homozygous to homozygous (either common or variant allele) switches (1.13%) and allele deletions or gains (0.5%).

Table 2. Overall genotyping concordance.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Agreement, %</th>
<th>κ-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>All markers</td>
<td>97.4</td>
<td>0.915</td>
</tr>
<tr>
<td>Polymorphic markers</td>
<td>94.3</td>
<td>0.812</td>
</tr>
</tbody>
</table>

Overall agreement and κ-statistic were calculated with all 1931 markers and with only the 870 polymorphic markers: markers for which the minor allele was detected in at least one sample.

Of all markers, those with the 20% lowest κ values (all κ<0.4) were considered poorly concordant, excluding markers with a low κ but agreement >80% because for those markers the κ value is strongly affected by a low minor allele frequency. Visual inspection revealed that only 9 of the 42 markers had a call rate <80% across all 32 samples, indicating that the marker concordance is not highly linked to the general marker call rate. For 28 of 42 markers, the cluster plots were considered unreliable (Supplementary Table S2).

In addition, we assessed whether the discordant results were randomly distributed by testing for associations of the 1,931 markers with DNA source (blood or saliva versus FFPE tissue), including genotype calls only. Thirteen markers had significant differences in minor allele frequencies between blood or saliva and FFPE DNA samples (P<0.05). Of these markers, nine were excluded because of unreliable cluster plots (CYP11B1 rs4534, CYP2A6 rs4079369, CYP2A7 rs4079366, CYP2B6 rs36079186, CYP2C9 rs1057910, CYP2D6 rs5030865, CYP2D6 rs61736512, SULT1A2 rs1059491, TBXAS1 rs8192868), leaving four SNPs (FMO6 rs2272797, CYP2B6 rs2273934, CYP39A1 rs2277119, and GSTM1 rs74837985), with only the latter (agreement 56.3%, κ=0.420) not overlapping with the top 42 markers with lowest concordance.

The 42 markers with the most discordant results also included markers located in genes for which the DMET Plus array performs automated haplotype calling and phenotype translation. Because calling and translation automatically involve all markers per gene, discordances in haplotype and subsequent phenotype results were observed in one or more sample pairs for CYP2B6, CYP2C9, CYP2D6, SLCOB1, TBXAS1, TPMT, and UGT1A1.
Discussion

There is a large worldwide collection of stored FFPE tissue, which can be highly valuable for providing germline DNA for genetic studies. In this study, we examined the utility of DNA isolated from normal FFPE tissue blocks for analysis on the DMET Plus array, which is a useful tool for pharmacogenetic studies. We found that the DNA isolated from FFPE tissue performs well on the array and yields high quality data. Not only did we observe high call rates, but importantly we also noted that the concordance with genotyping results of matching blood or saliva samples was high. These data indicate that DNA from FFPE tissue can be reliably used as a template for the DMET Plus array.

Factors related to tissue fixation and storage affect the quality of the DNA that can be recovered from the specimen. SNP arrays based on MIP technology may be an optimal solution for working with such degraded DNA because of the small target DNA sequences (minimum target length of 40 bp). A previous study by Wang et al. using an Affymetrix MIP array with a panel of 50,000 markers found a mean call rate of 98.9% for DNA from normal FFPE tissue samples, similar to the mean call rate observed in the current study [10]. Hence, our data for the DMET Plus array extend these observations, suggesting that even though the DMET procedure includes a multiplex PCR step, this MIP-based array is indeed suitable for the successful genotyping of DNA of reduced quality. The combination of the multiplex PCR step to preamplify regions in genes with pseudogenes and close homologs and the MIP technology has previously been found to yield reliable genotypes in these challenging genes (e.g. CYP2D6, UGT1A1) in a validation study of the DMET Plus array using blood samples [17]. Our study reveals that reliable genotypes can be obtained for markers of these challenging genes even in the context of FFPE-derived DNA.

In our study, we obtained reliable genotyping results from FFPE-derived DNA samples on the DMET Plus array, even for the samples with degraded DNA based on the size range PCR. Apparently, the primers for the DMET Plus array have been designed in such a way that they are very close to the SNP of interest. On the basis of our test panel, the prescreening of samples based on the size range PCR does not appear to be an essential control to determine whether samples are of sufficient quality for further DMET analysis. We observed two samples with relatively low call rates, which corresponded to a mild decrease in concordance compared to samples with call rates >98%. Therefore, we suggest evaluating the call rates of FFPE-derived DNA samples before further genetic analyses because we found that low call rates are related to a lower concordance between blood or saliva and FFPE tissue. Therefore, the call rates are related to the reliability of the data. In case call rates of a small number of samples are too low for reliable association analysis, these samples could still be included in a meta-analysis together with the DMET genotyped samples by specifically genotyping the significantly associated markers using other genotyping methods (e.g. TaqMan allelic discrimination assays) [11].

The analysis of the discordant results indicated that certain SNPs were nonrandomly miscalled. Despite the limited number of samples in the association analysis, it cannot be excluded that genotyping results for the four SNPs that were found to be associated with DNA source are systematically unreliable in FFPE DNA samples. Therefore, we suggest caution when including these SNPs in association analyses with the (clinical) outcome of interest. The same applies to the reported list of markers with poor concordance, since these markers are prone to discrepancies as well. However, as indicated, some of these markers show unreliable genotype clustering, which might not only explain the sensitivity to sample quality variation but might indicate the necessity to exclude them from further analyses regardless of the sample type used as DNA source. The VIP genes covered by the DMET Plus array had good overall concordance; for those VIP genes that had the least agreement, this could be explained by the presence of markers reported here to have poor concordance. As subsequently expected, the discordant markers were found to impact the haplotype calling and phenotype translation as performed by the DMET console software, which is explained by absent or incorrect genotype calls in the FFPE DNA samples. Hence, exclusion of the reported discordant markers is warranted in manual translation, but could still affect reliability because of the missing genotypes. However, of the alleles that had poor concordance and that are used in clinical guidelines, most are expected to have a low impact because of their low allele frequency.

We used only normal FFPE tissue sections to isolate germline DNA, avoiding different genotyping results in DNA from FFPE tissue compared with blood or saliva because of somatic alterations in tumor DNA. The issue of using tumor DNA in pharmacogenetic studies has been subject of debate, and although germline DNA is preferred, recent studies have provided evidence that DNA from FFPE tumor tissue blocks may be used as a valid alternative in pharmacogenetic cancer studies when no other samples are available [18,19]. However, in the current study, discordant results that may be produced when using FFPE tumor tissue on the DMET Plus array are not explored.

Our published optimized protocol for DNA isolation from FFPE tissue for TaqMan SNP assays could also optimally be used in the current study for the generation of samples suitable for downstream use on the DMET Plus array [11]. For single SNP assays, not only nondecalcified but also decalcified FFPE tissue has been shown to be useful. However, for the DMET analysis 1.02 µg of DNA (60 ng/µL) is needed per sample, which is often difficult to obtain from tissues decalcified with formic acid because this treatment makes the recovery of DNA even more challenging. Therefore, no samples from decalcified FFPE tissue could be included in the current study. A similar limitation of the use of FFPE-derived DNA for the DMET array may also arise in case researchers are restricted to tissues with low abundance of nuclei, from which the necessary quantity of DNA cannot...
be recovered. Whole genome amplification based on multiple displacement amplification has been reported to provide accurate results on the DMET Plus array and may provide a solution, although caution should be taken when using input material of low quality [20]. To date, there are no reports on the reliability of whole genome amplification of FFPE-derived DNA for use on the DMET Plus array.

In conclusion, we found that high-quality genotype data can be obtained using the DMET Plus array on DNA isolated from FFPE samples. This provides opportunities for pharmacogenetic studies to make use of archival tissue resources to analyze individuals not otherwise available for sampling.

References


<table>
<thead>
<tr>
<th>Gene</th>
<th>No. of markers on DMET (no. of polymorphic markers)</th>
<th>Agreement %</th>
<th>κ</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCB1</td>
<td>40 (17)</td>
<td>94.8</td>
<td>0.907</td>
</tr>
<tr>
<td>ADH1A</td>
<td>7 (3)</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>ADH1B</td>
<td>6 (3)</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>ADH1C</td>
<td>5 (1)</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>AHR</td>
<td>6 (3)</td>
<td>97.9</td>
<td>0.807</td>
</tr>
<tr>
<td>ADCY3A1</td>
<td>7 (3)</td>
<td>99.1</td>
<td>0.857</td>
</tr>
<tr>
<td>ADAT</td>
<td>5 (2)</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>CYP1A2</td>
<td>17 (4)</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>CYP2A6</td>
<td>19 (9)</td>
<td>92.4</td>
<td>0.736</td>
</tr>
<tr>
<td>CYP2B6</td>
<td>23 (14)</td>
<td>87.8</td>
<td>0.580</td>
</tr>
<tr>
<td>CYP2C9</td>
<td>18 (4)</td>
<td>95.3</td>
<td>0.935</td>
</tr>
<tr>
<td>CYP2CR</td>
<td>11 (4)</td>
<td>98.9</td>
<td>0.818</td>
</tr>
<tr>
<td>CYP2C9</td>
<td>18 (6)</td>
<td>95.8</td>
<td>0.796</td>
</tr>
<tr>
<td>CYP2D6</td>
<td>30 (16)</td>
<td>82.7</td>
<td>0.725</td>
</tr>
<tr>
<td>CYP2E1</td>
<td>10 (7)</td>
<td>98.1</td>
<td>0.830</td>
</tr>
<tr>
<td>CYP2J2</td>
<td>9 (2)</td>
<td>99.3</td>
<td>0.889</td>
</tr>
<tr>
<td>CYP3A4</td>
<td>25 (6)</td>
<td>99.3</td>
<td>0.946</td>
</tr>
<tr>
<td>CYP3A5</td>
<td>14 (3)</td>
<td>98.7</td>
<td>0.852</td>
</tr>
<tr>
<td>DPID</td>
<td>18 (6)</td>
<td>97.9</td>
<td>0.899</td>
</tr>
<tr>
<td>G6PD</td>
<td>6 (6)</td>
<td>99.0</td>
<td>0.979</td>
</tr>
<tr>
<td>GSTP1</td>
<td>5 (2)</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>GSTT1</td>
<td>6 (6)</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>HMGCR</td>
<td>6 (2)</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>ME2</td>
<td>16 (6)</td>
<td>98.0</td>
<td>0.912</td>
</tr>
<tr>
<td>NQO1</td>
<td>7 (4)</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>NQ1B2</td>
<td>15 (3)</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>PGM5</td>
<td>12 (1)</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>SCC19M1</td>
<td>5 (3)</td>
<td>68.8</td>
<td>0.526</td>
</tr>
<tr>
<td>SCC22A1</td>
<td>25 (11)</td>
<td>95.0</td>
<td>0.769</td>
</tr>
<tr>
<td>SCOT1B1</td>
<td>18 (8)</td>
<td>95.8</td>
<td>0.860</td>
</tr>
<tr>
<td>SUZ2A1</td>
<td>6 (1)</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>TPM7</td>
<td>8 (6)</td>
<td>87.5</td>
<td>0.382</td>
</tr>
<tr>
<td>TYMS</td>
<td>5 (0)</td>
<td>100</td>
<td>1.00</td>
</tr>
<tr>
<td>GSTT1A1</td>
<td>32 (10)</td>
<td>96.9</td>
<td>0.949</td>
</tr>
<tr>
<td>VKORC1</td>
<td>10 (10)</td>
<td>99.4</td>
<td>0.987</td>
</tr>
</tbody>
</table>

Overall agreement and κ-statistic calculated per VIP gene with all markers per gene included.
Genotyping FFPE tissue DNA on DMET array

---

**Supplementary Table S2, DMET Plus markers showing lowest concordance.**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2235022</td>
<td>ABCB1</td>
<td>31.3</td>
<td>0.000</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs28189115</td>
<td>ABCB1</td>
<td>0.0</td>
<td>0.000</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs4148807</td>
<td>ABCB4</td>
<td>56.3</td>
<td>0.000</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>rs17316324</td>
<td>ABC2</td>
<td>6.25</td>
<td>0.000</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs1437810</td>
<td>CHST2</td>
<td>68.8</td>
<td>0.000</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>rs1046349</td>
<td>CHST7</td>
<td>75.0</td>
<td>0.366</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>rs28693444</td>
<td>CHST9</td>
<td>75.0</td>
<td>0.000</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs5134</td>
<td>CYP1B1</td>
<td>12.5</td>
<td>0.034</td>
<td>11</td>
<td>7</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs72549371</td>
<td>CYP1B1</td>
<td>16.8</td>
<td>0.000</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>rs7252758</td>
<td>CYP21A2</td>
<td>12.5</td>
<td>0.000</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs7252257</td>
<td>CYP21A2</td>
<td>75.0</td>
<td>0.143</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>rs4079669</td>
<td>CYP2A6</td>
<td>18.8</td>
<td>0.000</td>
<td>13</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs4079566</td>
<td>CYP2A6</td>
<td>56.3</td>
<td>0.000</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>+</td>
</tr>
<tr>
<td>rs36097168</td>
<td>CYP2B1*13</td>
<td>75.0</td>
<td>0.000</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs2272797</td>
<td>CYP2B1*13</td>
<td>75.0</td>
<td>0.200</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs35546288</td>
<td>SLC22A1</td>
<td>75.0</td>
<td>0.000</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs55918055</td>
<td>SLC22A1</td>
<td>75.0</td>
<td>0.000</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs11231820</td>
<td>SLC22A1</td>
<td>68.8</td>
<td>0.000</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs4149015</td>
<td>SLCOB1*17</td>
<td>56.3</td>
<td>0.152</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs75547537</td>
<td>SULT1A1*4</td>
<td>18.8</td>
<td>0.000</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs1054981</td>
<td>SULT1A2</td>
<td>43.8</td>
<td>0.200</td>
<td>9</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs35044222</td>
<td>SULT1A3</td>
<td>0.0</td>
<td>0.004</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>

All markers with κ<0.4 and agreement <80%. Of the reported variants, Clinical Pharmacogenetics Implementation Consortium guidelines are available for the low frequency alleles TPMT*3B and the CYP2D6 alleles, as well as for CYP2C9*3 and SLCOB1*17.

PRA: possible rare allele; Hom: homozygous call; Het: heterozygous call.

---

**Supplementary Table S2, Continued**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>rs35046288</td>
<td>SLC22A1</td>
<td>75.0</td>
<td>0.000</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs55918055</td>
<td>SLC22A1</td>
<td>75.0</td>
<td>0.000</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs11231820</td>
<td>SLC22A1</td>
<td>68.8</td>
<td>0.000</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs4149015</td>
<td>SLCOB1*17</td>
<td>56.3</td>
<td>0.152</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs75547537</td>
<td>SULT1A1*4</td>
<td>18.8</td>
<td>0.000</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs1054981</td>
<td>SULT1A2</td>
<td>43.8</td>
<td>0.200</td>
<td>9</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs35044222</td>
<td>SULT1A3</td>
<td>0.0</td>
<td>0.004</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs35044222</td>
<td>SULT1A3</td>
<td>12.5</td>
<td>0.018</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs1912868</td>
<td>TDRD5*7</td>
<td>12.5</td>
<td>0.000</td>
<td>14</td>
<td>11</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs1200460</td>
<td>TPMT*3B</td>
<td>62.5</td>
<td>0.056</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>rs1424874</td>
<td>UGT1A1*40</td>
<td>62.5</td>
<td>0.356</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>rs10409453</td>
<td>VKORC1</td>
<td>18.8</td>
<td>0.056</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>rs13336604</td>
<td>VKORC1</td>
<td>75.0</td>
<td>0.103</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>
Adverse Drug Reactions
Chapter 3

Replication of a genetic variant in ACYP2 associated with cisplatin-induced hearing loss in patients with osteosarcoma


* These authors contributed equally

Pharmacogenetics and Genomics 2016 May;26(5):243-7
Abstract

Objective: Irreversible hearing loss is a frequent side effect of the chemotherapeutic agent cisplatin and shows great interpatient variability. The variant rs1872328 in the ACYP2 gene was recently identified as a risk factor for the development of cisplatin-induced ototoxicity in children with brain tumors. We aimed to replicate this finding in patients with osteosarcoma.

Methods: An independent cohort of 156 patients was genotyped for the rs1872328 variant and evaluated for the presence of cisplatin-induced ototoxicity.

Results: A significant association was observed between carriership of the A allele and cisplatin-induced ototoxicity after end of treatment ($P=0.027$).

Conclusion: This is the first study replicating the association of ACYP2 variant rs1872328 with cisplatin-induced ototoxicity in patients with osteosarcoma who did not receive potentially ototoxic cranial irradiation. Hence, the ACYP2 variant should be considered a predictive pharmacogenetic marker for hearing loss, which may be used to guide therapies for patients treated with cisplatin.

Introduction

Cisplatin is an important component of the chemotherapeutic treatment of a variety of solid tumors, including osteosarcoma. A frequent side effect of cisplatin treatment is ototoxicity, which occurs in 40-60% of the patients [1]. Ototoxicity is characterized by irreversible, bilateral, sensorineural hearing loss, which is a serious complication especially in the pediatric population because of its impact on speech and language development. Moreover, ototoxicity is one of the main reasons for dose reductions or even termination of cisplatin treatment.

Several clinical risk factors for cisplatin-induced ototoxicity have been described, including young age at diagnosis (<5), cumulative cisplatin dose, use of additional ototoxic drugs such as carboplatin, and cranial irradiation [2-5]. However, these parameters are not sufficient to reliably predict cisplatin-induced ototoxicity. To tailor the treatment for the individual patient and avoid ototoxicity, there is a need to predict this side effect at the start of treatment. This would enable treatment optimization to minimize the risk to susceptible groups, for example, by using protective agents.

A number of publications indicate that genetic factors might be useful predictors; however, most of these studies involve a limited number of patients without replication in a second population [6-9]. In addition, the most promising candidate gene study published to date, suggesting that genetic variants in the COMT and TPMT genes are linked to cisplatin-induced ototoxicity, could not be consistently replicated by ourselves and others [10-13].

Recently, Xu et al. identified a genetic variant (rs1872328) in ACYP2 that predisposes to precipitous hearing loss in children with newly diagnosed embryonal brain tumors, using a genome-wide association approach including 238 patients and a replication cohort of 68 patients [14]. The ACYP2 gene encodes an acylphosphatase expressed in muscle and the cochlea and may be involved in hair cell development [14,15]. The variant in this gene might also be related to cisplatin-induced hearing loss in other diseases where slightly different treatment regimens are used.

Patients with osteosarcoma are treated with a high cumulative dose of cisplatin and, in addition, no (cranial) irradiation is applied, which is a confounding factor for cisplatin-induced hearing loss. Hence, this patient population represents an excellent group to study the generalizability of the ACYP2 gene variant to predict cisplatin-induced ototoxicity. Therefore, we investigated an independent osteosarcoma patient cohort and carried out a meta-analysis including the previously published study [14], resulting in a total population of 462 patients.
Methods

Patient cohort
A cohort of Dutch patients with primary, high-grade osteosarcoma treated with cisplatin was retrospectively recruited at the Radboud university medical center (Nijmegen), the University Medical Center of Groningen (Groningen), Leiden University Medical Center (Leiden) and the Emma Children’s Hospital/Academic Medical Center (Amsterdam), The Netherlands. Eligibility also included a maximum age at diagnosis of 45 years, that is, only patients from the first osteosarcoma incidence peak were included. From the total cohort, those patients were included for whom audiometric data were available from at least end of treatment (N=156, all treated between 1984 and 2013). All patients were of self-reported European descent, except for one patient originating from Brazil. Data on co-administration of other potentially ototoxic drugs such as carboplatin and vincristine were recorded [16,17]. None of the patients was treated with cranial irradiation or otoprotective agents. Germline DNA was extracted from blood or saliva using the QIAamp DNA Blood Midi kit (Qiagen, Venlo, The Netherlands) and collection kits from DNA Genotek (Kanata, Ontario, Canada), respectively. From patients who had died before inclusion, germline DNA was isolated from paraffin-embedded tissue samples as described previously [18]. The study protocol was approved by the Institutional review board of the Radboud university medical center (Commissie Mensgebonden Onderzoek Regio Arnhem Nijmegen) and approval for inclusion of patients in other institutes was obtained from institutional ethics committees; all patients and/or parents provided written informed consent.

Ototoxicity
All audiometric assessments were age appropriate and performed by conventional or play audiometry under standardized conditions as part of routine clinical monitoring for cisplatin-related hearing loss.

Grading was based on the ear with the worst hearing as assessed by the latest measurement available, that is, the exact follow-up time is different for each patient. Hearing loss was retrospectively classified according to the Chang grading system [19]. In case of known hearing loss at baseline, ototoxicity was graded based on the change in the Chang score as suggested by Chang and colleagues.

Genotyping
Genotyping for the ACYP2 variant rs1872328 was performed using a Kaspar-On-Demand (KOD) assay (LGC Genomics, Hoddesdon, UK). Genotyping was carried out in a volume of 5 µL containing 10 ng of genomic DNA, 2.5 µL of KASP 5000 V4.0 High ROX (2x; LGC Genomics) and 0.0625 µL of the KASPar assay (40x), and 1.44 µL of MilliQ grade water. Each amplification for the KASPar assay was performed by an initial denaturation at 94°C for 15 min, followed by 10 cycles of denaturation at 94°C for 20 s and annealing/extension at 61°C for 60 s including a drop of 0.6°C for each cycle. This was followed by 26 cycles of denaturation at 94°C for 10 s and annealing/extension at 55°C for 60 s, followed by 4 cycles of denaturation at 94°C for 20 s and annealing/extension at 57°C for 60 s. This was carried out on a 7500FAST Real-Time PCR System (Applied Biosystems, Nieuwerkerk aan den IJssel, The Netherlands). Genotypes were scored using the algorithm and software (v2.0.6) supplied by Applied Biosystems. Nontemplate controls (3%) as well as duplicates between plates were included as quality controls for genotyping.

Statistical analysis
Statistical differences in demographic data between patients who had hearing loss (Chang score > 0) and patients without hearing loss (Chang score = 0) were assessed using the Fisher exact, Pearson X²-test, or Mann-Whitney U-test as appropriate. Reported P-values are two-sided and are considered statistically significant if less than 0.05. To assess the effect of the genotype on susceptibility to ototoxicity, the data were dichotomized into score 0 versus scores 1-4. The association between genotype (A carriers vs. homozygous G) and ototoxicity was assessed using a 2x2 table (Fisher exact). All analyses were carried out using SPSS package, version 20 (SPSS Inc., Chicago, Illinois, USA). Meta-analysis of the data of ototoxicity after the end of treatment (Chang score >0), kindly provided by the authors of the original genome-wide association study (GWAS) publication, was carried out using a fixed-effects model in Review Manager, 5.3 (The Cochrane Collaboration, Oxford, UK). On the basis of the discovery cohort of Xu and colleagues with a minor allele frequency of 0.046 and trait prevalence of 0.61, a sample size of 156 patients was calculated to have 80% power to detect significant associations (P<0.05) with odds ratios of 7.5 and higher. For this, the genetic power calculator Quanto (University of Southern California, California, USA) was used [20].

Results
Of a total of 156 patients included, 77 (49.4%) presented with ototoxicity (Chang score >0) and 36 (23.1%) patients showed clinically relevant ototoxicity (Chang score ≥2a). No significant differences in patient characteristics were observed between patients with hearing loss (Chang score >0) and those without hearing loss (score=0) (Table 1). Therefore, no clinical covariates are included in the genetic analysis. A total of 12 patients were additionally treated with other ototoxic chemotherapeutic agents (Table 1). In 22 (14.1%) patients, the (cumulative) cisplatin dose was reduced because of ototoxicity.
Table 1. Patient demographics of the osteosarcoma cohort.

<table>
<thead>
<tr>
<th></th>
<th>Chang score = 0 (N=79)</th>
<th>Chang score &gt; 0 (N=77)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis* (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>14.2 (5.1-42.4)</td>
<td>14.0 (3.4-43.9)</td>
<td>0.94</td>
</tr>
<tr>
<td>Sex (male)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[n (%)]</td>
<td>39 (49.4)</td>
<td>45 (58.4)</td>
<td>0.26</td>
</tr>
<tr>
<td>Cumulative cisplatin dose (mg/m²)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>480 (200-600)</td>
<td>480 (140-720)</td>
<td>0.07†</td>
</tr>
<tr>
<td>Concomitant drugs [n (%)])</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vincristine</td>
<td>2 (2.5)</td>
<td>7 (9.1)</td>
<td>0.10</td>
</tr>
<tr>
<td>Carboplatin</td>
<td>2 (2.5)</td>
<td>1 (1.3)</td>
<td>1.0</td>
</tr>
<tr>
<td>Follow-up (days)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>666 (17-7,061)</td>
<td>830 (1-9,344)</td>
<td>0.91</td>
</tr>
</tbody>
</table>

*Patients ≥18 years (N=33): Chang score = 0 N=12; Chang score > 0 N=21.
†No trend is observed on exclusion of patients with dose reductions because of ototoxicity; P=0.429; therefore, cisplatin dose is not included as a covariate.

Table 2. Genotype frequencies on the basis of ototoxicity severity.

<table>
<thead>
<tr>
<th></th>
<th>GG genotype [n (%)]</th>
<th>A carriers [n (%)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chang score = 0 (N=79)</td>
<td>79 (100)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Chang score = 1 (N=40)</td>
<td>37 (92.5)</td>
<td>3 (7.5)</td>
</tr>
<tr>
<td>Chang score ≥ 2a (N=36)</td>
<td>34 (94.4)</td>
<td>2 (5.6)</td>
</tr>
</tbody>
</table>

Figure 1. Forest plot of meta-analysis of ACYP2 variant rs1872328. Meta-analysis of published cohorts and present study using a fixed-effects model. For effect size estimation, Review Manager has corrected for the zero-cell counts by adding a fixed value of 0.5 to all cells of the study results table. 95% CI, 95% confidence interval of odds ratio; M-H, Mantel-Haenszel method.

Discussion

Permanent hearing loss induced by cisplatin is a serious complication in the treatment of patients with solid tumors including osteosarcoma. In a recent GWAS in children with embryonal brain tumors, a variant in the ACYP2 gene was identified as a pharmacogenetic marker for cisplatin-induced ototoxicity [14]. In the present study, we have replicated the previously reported association of the ACYP2 variant with cisplatin-induced ototoxicity in an independent cohort of patients with osteosarcoma and additionally validated the effect in a meta-analysis with the cohorts from the previously published study. Not only does this and in 24 (15.4%) patients because of other cisplatin-induced toxicities. The cumulative cisplatin dose did not significantly differ between patients with and without hearing loss (Table 1), also after exclusion of patients with dose reductions because of ototoxicity (P=0.429). Baseline audiograms were missing for 54 (34.6%) patients, of whom all except one were aged less than or equal to 40 years and therefore not expected to present with age-related baseline hearing loss. From 102 patients with baseline measurements, six patients showed hearing loss before the start of cisplatin treatment. The distribution of the follow-up times until the last ototoxicity evaluation was comparable between patients with and without ototoxicity (Supplementary Figure S1). In patients without ototoxicity, similar to patients with ototoxicity, the majority of patients evaluated within 2 years after the end of cisplatin therapy survived more than 2 years after treatment, that is, detection of potential ototoxicity in this group was not limited by early follow-up because of short survival [21].

Five patients were heterozygous AG for the ACYP2 variant rs1872328; all other patients were homozygous GG. Patients carrying the A allele represented 6.5% of the patients with cisplatin-induced ototoxicity and 3.2% of all patients, corresponding to a minor allele frequency of 0.016. No deviation from Hardy-Weinberg equilibrium was observed (P=0.78). A statistically significant association was observed between carriership of the A allele and cisplatin-induced ototoxicity: 72 (47.7%) patients with the homozygous GG genotype developed cisplatin-induced ototoxicity, whereas ototoxicity was observed in all five patients with the AG genotype (P=0.027). After grouping of patients on the basis of the severity of ototoxicity, the association was still significant (P=0.033), although we did not observe an increase in the number of patients carrying the A allele with increased ototoxicity severity (Table 2).

A meta-analysis of the present study and the cohorts of the previously published study [14], including a total of 462 patients, also indicated a significant association of the ACYP2 variant with ototoxicity, with a pooled odds ratio of 14.7 (P=0.001) (Figure 1).
confirm the association but it also indicates that the association applies to different types of cancer with different treatment protocols, for example, in patients who did not receive cranial irradiation. Although there are still patients without the variant who presumably develop ototoxicity, the results have implications for the proportion of patients who carry this variant. The ACYP2 gene may be a useful marker to screen patients at diagnosis to enable informed clinical decisions on cisplatin treatment for patients carrying the A allele, including dose adjustments, alternative strategies, or use of protective agents.

Replication of pharmacogenetic findings is an essential step towards clinical implementation, but is frequently not achieved [22], which could be attributed to false-positive findings in the original study or because of reasons such as differences in treatment protocols and phenotyping, biased or heterogeneous patient populations, and insufficient sample sizes. Especially in pediatric solid tumor studies, patient numbers can be limited because of low disease prevalence. In the present study, we could replicate the finding by Xu and colleagues, even though our osteosarcoma cohort was treated with a higher cumulative cisplatin dose compared with the brain tumor study. Furthermore, we have studied the largest osteosarcoma patient cohort for pharmacogenetic research to date, which, despite its relatively small size, is sufficiently large to detect significant genetic associations with large effect sizes. The fact that we did detect the ACYP2 association in this relatively small cohort indicates that the ACYP2 variant is a genetic predictor with a large effect size.

Although multiple variants in the ACYP2 gene were potentially associated in the study by Xu et al. [14], we focused on the only genome-wide significant top hit, which was also the only variant reported to be replicated within the original study. Now that our study confirms the involvement of the ACYP2 gene in cisplatin-induced ototoxicity, it would be interesting to investigate other variants within the same region. Sequencing the gene would be an elegant approach, which was not performed in the present study due to limited sample size. This could indicate additional common variants but also rare variants with potentially larger effect sizes. This would enhance the predictability of ototoxicity on the basis of genetic variation in ACYP2; however, this requires large scale studies.

In the GWAS by Xu and colleagues, the ACYP2 variant was detected in a regression model treating ototoxicity as a time-dependent variable. In the present study, only the ototoxicity after ending treatment was analyzed as we could not carry out the time-to-event analysis because of the retrospective data collection. However, the association could still be confirmed when considering ototoxicity at end of treatment as the phenotype, which is still highly relevant for clinical practice as it shows that ACYP2 can be used to identify patients at risk of developing cisplatin-induced ototoxicity irrespective of time.

In previous studies, younger age, higher cumulative cisplatin doses, concomitant ototoxic drugs, and cranial irradiation have been reported as risk factors for cisplatin-induced ototoxicity. However, in our osteosarcoma cohort, patients did not receive cranial irradiation and none of the other clinical variables was found to be a predictor of ototoxicity. In clinical practice, in specific patients groups in which these factors are present and may vary significantly, the predictive value of these factors might be considered in combination with the genetic risk.

In our patient cohort as well as the two cohorts from the original study, the ACYP2 variant allele was only present in patients presenting with ototoxicity. The resulting zero-cell count in the groups without ototoxicity carrying the A allele affected the meta-analysis of all three studies because of difficulties in estimating odds ratios. However, the specialized meta-analysis software Review Manager corrected for this and thereby computational problems could be avoided.

In conclusion, we have shown that the previously reported association of an ACYP2 variant with cisplatin-induced hearing loss in pediatric brain tumor patients could be replicated in an independent cohort of patients with osteosarcoma who did not receive potentially ototoxic cranial irradiation. Further characterization of the molecular mechanism underlying this association and prospective studies to validate the clinical utility are needed. The ACYP2 variant should be considered a predictor of cisplatin-induced hearing loss in patients with osteosarcoma or other solid tumors and may contribute toward personalized treatment.
References


Figure S1. Distribution of follow-up time from end of cisplatin therapy to last ototoxicity evaluation. Histogram plots are given for both patients without ototoxicity (A) and with ototoxicity (B), and include indication of survival after end of cisplatin therapy based on a cut-off of 2 years.
Chapter 4

The role of genetic variants in ADME genes in high-dose methotrexate-induced hepatotoxicity in patients with osteosarcoma


* These authors contributed equally

Submitted
Abstract

Hepatotoxicity is a frequent and serious side effect in high-dose methotrexate (HD-MTX) treatment. We explored the contribution of 1,936 genetic variants in 231 drug absorption, distribution, metabolism and elimination (ADME) genes to interindividual variation in MTX plasma levels and hepatotoxicity in HD-MTX-treated osteosarcoma patients.

An osteosarcoma cohort (N=114) was genotyped using the DMET array. Genetic variants significantly associated with MTX plasma levels at 48 hours after MTX infusion were assessed for association with MTX-induced hepatotoxicity (CTCAE grade 3/4) by logistic regression.

Upon analysis of 1,250 HD-MTX courses, 32 genetic variants were associated with 48 hour MTX plasma levels. Of these variants, GSMT3 rs1799735 was associated with frequent hepatotoxicity (protective effect: OR 0.13 [0.02-0.73], P=0.021).

This is the first study in osteosarcoma patients in which a broad range of ADME genes was explored in relation to MTX plasma levels and HD-MTX-induced hepatotoxicity, indicating a role for a genetic variant in GSTM3.

Introduction

Methotrexate (MTX) is an antifolate agent widely used in the treatment of cancer. In high doses (≥1 g/m²), it is an important component of many chemotherapeutic protocols in pediatric malignancies, including osteosarcoma [1]. This bone tumor constitutes approximately 3% of all childhood cancers and is, despite its rarity, the most common primary malignancy of the bone in children and adolescents [2,3]. Together with cisplatin and doxorubicin, high-dose methotrexate (HD-MTX) forms the backbone of the treatment of osteosarcoma.

Hepatotoxicity is one of the most frequent side effects of HD-MTX and is a reason for delayed continuation of the chemotherapeutic treatment, which is inversely related to survival [4]. As exposure to high MTX levels is often related to hepatotoxicity, therapeutic monitoring of MTX plasma levels is employed during the first days after infusion, together with appropriate leucovorin rescue and intensive hydration to minimize HD-MTX-induced side effects. However, hepatotoxicity is still commonly observed and prolonged leucovorin rescue in case of high MTX plasma levels might reduce the desired antitumor effect of HD-MTX, leading to decreased treatment efficacy. A better understanding of the factors that contribute to interindividual pharmacokinetic and pharmacodynamic variability may allow upfront identification of patients at risk of toxic MTX exposure and thereby prevention of HD-MTX-induced hepatotoxicity, which may involve dose reduction although this is not without the concern of reducing efficacy as well.

Clinical risk factors do not sufficiently explain interindividual differences in response to HD-MTX, both regarding MTX-levels and toxicities, and several studies have suggested a role of genetic variation [5]. A genome-wide association study (GWAS) in patients with acute lymphoblastic leukemia (ALL), has identified variants in the organic anion transporter gene SLCO1B1 to be associated with MTX clearance [6,7]. For osteosarcoma patients, treated with even higher cumulative doses than ALL patients, a few small pharmacogenetic studies have been reported. A Slovenian study assessing tumor DNA of 44 osteosarcoma patients, has also suggested a role of SLCO1B1 in MTX clearance, and of ABCC2 in hepatotoxicity [8]. Contradicting results have been reported on associations of an MTHFR genetic variant and hepatotoxicity [9-11]. All these studies were based on a small number of candidate genes, and more comprehensive investigation of drug metabolism and transporter genes is warranted. Therefore, in the present study the Affymetrix Drug Metabolizing Enzymes and Transporters (DMET) Plus genotyping platform was used, which covers 1,936 variants in 231 genes involved in drug absorption, distribution, metabolism and elimination (ADME) [12]. The aim of this exploratory study was to identify genetic variants associated with 48 hour MTX plasma levels and HD-MTX-induced hepatotoxicity in HD-MTX-treated patients with osteosarcoma.
Materials and methods

Patients and treatment
A cohort of 114 patients with primary, high-grade osteosarcoma was retrospectively collected at the Radboud university medical center (Nijmegen), the University Medical Center of Groningen (Groningen), Leiden University Medical Center (Leiden) and the Emma Children’s Hospital/Academic Medical Center (Amsterdam), The Netherlands. All patients were treated between 2003 and 2014 and clinical data were retrospectively collected from medical records. Inclusion criteria for this study were: age ≤45 years, self-reported Caucasian ethnicity, and treatment according to the EURAMOS-1 protocol (also including patients diagnosed outside of the trial recruitment period, but treated according to the same scheme) [13]. The study protocol was approved by the institutional review board of the Radboud university medical center (Commissie Mensgebonden Onderzoek Regio Arnhem Nijmegen), and approval for inclusion of patients in other institutes was obtained from institutional ethics committees. All patients and/or parents provided written informed consent.

All patients received a maximum of 12 courses HD-MTX (12 g/m² per course) as a 4 hour infusion. Leucovorin rescue (for all patients 15 mg/m² according to protocol) was started 24-28 hours after start of MTX infusion and was prolonged if MTX plasma levels were >0.40 µmol/L at 48 hours. In addition to MTX, all patients received doxorubicin (maximum cumulative dose: 450 mg/m²) and cisplatin (maximum cumulative dose: 480 mg/m²), with or without additional ifosfamide/etoposide or interferon-α.

Methotrexate plasma levels and toxicity
MTX plasma levels were routinely monitored by a fluorescence polarization immunoassay (TDx/FLx, Abbott Diagnostics, The Hague, The Netherlands), or enzyme immunoassay (for 3 of 114 patients) (Syva Emit TDM assay, Siemens Healthcare, Hoofddorp, The Netherlands), without differences in reference values. The MTX plasma levels measured >0.40 µmol/L, at 48 hours. In addition to MTX, all patients received doxorubicin (maximum cumulative dose: 450 mg/m²) and cisplatin (maximum cumulative dose: 480 mg/m²), with or without additional ifosfamide/etoposide or interferon-α.

DNA isolation
Germline DNA was isolated from blood (N=54) using the QIAamp DNA Blood Midi kit (Qiagen, Venlo, The Netherlands), or from saliva (N=53) using the Oragene saliva collection kit (DNA Genotek, Kanata, Ontario, Canada) according to the manufacturer’s protocols. From patients who had passed away before inclusion (N=7), DNA was isolated from normal formalin-fixed, paraffin-embedded (FFPE) tissue as described previously [14,15].

Genotyping
The DNA samples were genotyped for 1,936 genetic variants using the Affymetrix DMET Plus array according to the manufacturer’s instructions (Affymetrix UK Ltd, High Wycombe, UK). All samples were normalized to 60 ng/µL based on NanoDrop measurements (NanoDrop 2000; Thermo Scientific, Breda, The Netherlands). Genotypes were calculated with DMET console software 1.3 using the Dynamic Genotype Boundaries version 2 algorithm. Variants were excluded from analysis if the genotype cluster plots were considered unreliable, being plots with genotype calling showing merged clusters without distinct cluster boundaries. Additional stringent evaluation of the genotype clustering in combination with expected genotype frequencies was carried out for variants significant in association analysis. Samples and variants were excluded if call rates <0.9, minor allele frequency (MAF) <0.01 and/or deviating from Hardy-Weinberg equilibrium (HWE) (P-value <0.0001). Five copy number variants, 46 X-chromosomal variants, and one tri-allelic variant present on the array were not included in the analyses. Quality control was based on genotypes of a larger previously described osteosarcoma cohort (N=316), from which the 114 patients included in the present study were selected [16].

Statistical analysis
High MTX levels are linked with the development of MTX-induced toxicities [17]. Therefore, first the association analyses between genetic variants and MTX plasma levels were performed to select genetic variants that were likely to be the most relevant for the development of hepatotoxicity, based on the assumed mechanism.

Associations of genotypes with 48h MTX plasma levels were assessed by Generalized Estimating Equation (GEE) analysis using STATA (Stata Corporation, College Station, TX, USA). This is an extended linear regression model that takes into account the correlation between the repeated measurements within individual patients. Analyses were performed both treating the MTX levels as a continuous variable (ln transformed plasma levels) and as a dichotomized categorical variable (based on cut-off at 0.40 µmol/L at 48h). Regression coefficients/odds ratios (OR) and corresponding 95% confidence intervals were generated for each variant under the assumption of an additive model. For analysis, genotypes were coded as 0 [homozygous for major allele], 1 [heterozygous] or 2 [homozygous for minor allele].

Genetic variants associated with 48h MTX plasma levels were selected for association analysis with hepatotoxicity as outcome, using logistic regression analysis in PLINK (v1.07)[18]. Hepatotoxicity was defined as elevated ALAT and/or ASAT levels after one or more MTX infusions, on the basis of dichotomized CTCAE grades: grade 0, 1 versus
2 [no toxicity] and grade 3 and 4 [toxicity]. Patients without recorded ALAT levels for any of the MTX infusions were excluded. In addition, patients with frequent hepatotoxicity (elevated ALAT and/or ASAT levels after ≥4 of the MTX infusions) were compared to patients without hepatotoxicity. Both analyses were also performed with the patients showing grade 2 ALAT and/or ASAT levels included as cases instead of controls, as this is clinically considered as moderate hepatotoxicity, though without consequences for further treatment. Potential associations of age at diagnosis or sex with 48h MTX plasma levels were tested using linear and logistic GEE models (with MTX levels as continuous or dichotomized variable). Associations of sex, age at diagnosis or cumulative MTX dose with hepatotoxicity were assessed by the Pearson chi-square or Mann-Whitney U test as appropriate using IBM SPSS Statistics version 20 (SPSS Inc., Chicago, Ill, USA).

Additionally, association of the MTX plasma levels with 5-year progression-free survival (PFS) (time interval from diagnosis to progression or recurrence) was assessed by Cox proportional hazards models in SPSS, with either all repeated measurements as independent variable or the mean plasma level for each patient. All reported P-values are two-sided and were considered statistically significant if <0.05. No correction for multiple testing was performed because of the exploratory character of the study.

Results

Patient characteristics
The 114 included patients with osteosarcoma received a total of 1,368 MTX courses, with a median of 12 courses per patient (range 2-12) (Table 1). Reasons for cancelled courses included poor physical conditions of the patient (whether or not caused by the treatment), a patient’s own request, or death. MTX plasma level data at 48 hours was available for 1,250 courses, for which in 264 (21.1%) courses MTX plasma levels >0.40 µmol/L were observed. Neither sex nor age at diagnosis was significantly associated with the MTX plasma levels (both as continuous or dichotomized parameter). The 48h MTX plasma levels were not associated with 5-year PFS (P=0.926 both with the mean plasma level or all measured plasma levels as covariable). Eighty-three patients developed hepatotoxicity (grade 3 or 4 of ALAT and/or ASAT) after at least one HD-MTX course; of all patients with elevated ASAT levels, also elevated ALAT levels were observed after one of the courses. Hepatotoxicity (grade 3 or 4 of ALAT and/or ASAT) was observed in 259 of 800 courses for which data were available. In 74 of the 259 courses (28.6%) with hepatotoxicity, MTX plasma levels >0.40 µmol/L were observed. Age at diagnosis was significantly associated with hepatotoxicity (P=0.022) and frequent hepatotoxicity (P= 0.000052) (Table 2).

Genotyping
All 114 DNA samples were successfully genotyped using the DMET Plus array. Of the 1,884 variants included in the quality control, 90 variants were excluded based on unreliable clusterplots, 28 variants were excluded because of genotype call rates of <0.9 and 1,056 variants because of a MAF of <0.01. All variants left were in HWE. After quality control, 710 variants were left for analysis.

Genetic variants associated with MTX plasma levels and hepatotoxicity
In univariable GEE association analyses, 69 variants were significantly associated with 48 hour MTX plasma levels (continuous variable) and 58 variants with high 48 hour MTX plasma levels (dichotomized based on cut-off at 0.40 µmol/L), of which 35 variants were significantly associated in both analyses. Three of the 35 variants, ALDH3A2...
rs72547554, PPARD rs7771474, and POR rs2286824, were excluded after additional stringent evaluation of the genotype clustering, leading to 32 remaining variants in 22 genes (Table 3). Only those variants that were not only associated with the plasma levels as a continuous variable but also associated with clinically high MTX levels (in logistic regression analysis with levels ≥0.40 µmol/L), were selected for association analysis with hepatotoxicity, to increase the chance of testing the most relevant variants in relation to hepatotoxicity.

In multivariable analysis including age at diagnosis, none of the variants showed a significant association with hepatotoxicity, but when only patients with frequent hepatotoxicity (i.e. after ≥4 of 12 courses) were included, the Glutathione S-Transferase Mu 3 (GSTM3) variant rs1799735 (delAGG) (GSTM3*9) was significantly associated with frequent hepatotoxicity (grade 2) as cases did not reveal significant associations indicating a protective effect of the deletion. Association analyses also including patients with moderate hepatotoxicity (grade 2) showed an association with frequent hepatotoxicity, to increase the chance of testing the most relevant variants in relation to hepatotoxicity.

In table 3, Genetic variants significantly associated with 48 hour MTX plasma levels.

| Gene | SNP Chr Minor allele MAF OR 95% CI P P | MTX levels dichotomized | MTX levels continuous |
|------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| EPHX1 | rs2234698 1 C 0.018 11.01 4.33 27.96 <0.001 | 0.88 0.42 1.34 <0.001 |
| CYP2B6 | rs34223104 19 C 0.018 3.81 2.03 7.14 <0.001 | 0.48 0.27 0.68 <0.001 |
| CYP2C8 | rs1059930 10 G 0.057 3.08 1.76 5.39 <0.001 | 0.42 0.11 0.73 0.009 |
| CYP2A11 | rs1056937 10 T 0.420 0.57 0.41 0.81 0.002 | -0.15 -0.25 -0.04 0.006 |
| SLC6A6 | rs2341970 3 T 0.252 0.48 0.30 0.76 0.002 | -0.19 -0.31 -0.07 0.002 |
| CYP2B1 | rs1056836 2 G 0.415 0.58 0.41 0.83 0.002 | -0.14 -0.25 -0.04 0.008 |
| ADPR | rs1801244 13 C 0.461 1.78 1.20 2.64 0.004 | 0.15 0.03 0.26 0.012 |
| ADPRB | rs1801243 13 G 0.469 1.76 1.19 2.60 0.005 | 0.15 0.03 0.26 0.012 |
| CHST11 | rs2029885 13 A 0.470 0.49 0.30 0.76 0.006 | -0.29 -0.46 -0.12 0.001 |
| ADPRB | rs1051332 13 A 0.452 1.74 1.17 2.59 0.006 | 0.14 0.02 0.25 0.004 |
| CYP2A6 | rs592946 19 A 0.413 1.00 0.53 0.94 0.007 | -0.27 -0.49 -0.05 0.019 |
| CYP2AF1A8 | rs17763762 2 A 0.409 0.71 0.54 0.92 0.011 | -0.21 -0.35 -0.07 0.003 |
| ADC1B8 | rs1737413 4 T 0.009 0.55 0.34 0.86 0.012 | -0.70 -1.01 -0.39 0.000 |
| CYP2AF1 | rs7670668 19 C 0.004 0.71 0.55 0.93 0.012 | 0.18 0.10 0.26 0.000 |
| CYP2AF1A6 | rs17803783 2 T 0.035 0.33 0.13 0.83 0.019 | -0.27 -0.51 -0.03 0.025 |
| ADPR | rs7334118 13 G 0.022 0.24 0.08 0.80 0.020 | -0.13 -0.25 0.00 0.043 |

Table 3 continued.

| Gene | SNP Chr Minor allele MAF OR 95% CI P P | MTX levels dichotomized | MTX levels continuous |
|------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| PPARD | rs457813 6 T 0.026 0.33 0.13 0.84 0.020 | -0.26 -0.48 -0.08 0.007 |
| GSTB | rs141254 4 T 0.022 0.25 0.08 0.80 0.020 | -0.43 -0.70 -0.16 0.002 |
| GSTF11 | rs3736599 4 A 0.083 0.50 0.28 0.91 0.023 | -0.26 -0.40 -0.12 0.000 |
| CHST10 | rs3749910 2 C 0.469 0.67 0.46 0.97 0.032 | -0.13 -0.23 -0.02 0.018 |
| SLC2A4 | rs34139495 6 A 0.018 0.15 0.02 0.90 0.038 | -0.18 -0.32 -0.04 0.012 |
| ABC4 | rs1189466 13 T 0.053 0.37 0.14 0.95 0.038 | -0.21 -0.41 0.00 0.046 |
| ABC4 | rs1678319 13 A 0.053 0.37 0.14 0.95 0.038 | -0.21 -0.41 0.00 0.046 |
| CYPA4 | rs709882 19 G 0.070 1.90 1.03 3.52 0.040 | 0.27 0.02 0.52 0.033 |
| CHST10 | rs1530031 2 A 0.467 0.68 0.46 0.98 0.041 | -0.12 -0.23 -0.02 0.020 |
| GSTM3 | rs1799735 1 delAGG 0.118 0.54 0.29 0.98 0.044 | -0.17 -0.31 -0.03 0.015 |
| SLC2A4 | rs140718 3 G 0.456 1.42 1.01 2.01 0.045 | 0.11 0.01 0.21 0.029 |
| SLC2A4 | rs173248 3 C 0.456 1.42 1.01 2.01 0.045 | 0.11 0.01 0.21 0.029 |
| ABC4 | rs185574 3 C 0.456 1.42 1.01 2.01 0.045 | 0.11 0.01 0.21 0.029 |
| PPARD | rs1051741 1 T 0.110 0.52 0.27 0.99 0.045 | -0.14 -0.32 0.00 0.045 |
| CHST10 | rs1828193 2 C 0.478 0.72 0.51 1.00 0.048 | -0.11 -0.20 -0.02 0.020 |

MAF, minor allele frequency; OR, odds ratio; 95% CI, 95% confidence interval. Effect sizes and 95% CI are reported for the minor allele, an OR<1 and coefficient<0 indicate a protective effect of the minor allele. Variants within CHST10, SLC2A4, CYP1B1, ATP7B (except for rs7334118), and ABC4 are in linkage disequilibrium with r²>0.8.

**Discussion**

The candidate gene approach studies of MTX in osteosarcoma patients have previously shown encouraging results; however, most studies focused on functional genetic variants involved in MTX metabolism. In this study, we have looked beyond these candidate genes to explore associations of genes that are not known to be related to MTX pharmacokinetics. By analyzing a large number of variants in genes coding for drug metabolizing enzymes and transporters, genetic variants in 22 genes were found to be associated with high MTX plasma levels, of which the intronic GSTM3 variant rs1799735 showed an association with frequent hepatotoxicity.
The present study is the first to report an association of GSTM3 with MTX plasma levels and hepatotoxicity. The gene encodes a glutathione S-transferase which is known for detoxification of electrophilic compounds, including carcinogens, therapeutic drugs, environmental toxins and products of oxidative stress, by conjugation with glutathione [19,20]. Another member of the GST family, Glutathione S-Transferase Mu 1 (GSTM1), has previously been implicated in drug-induced hepatotoxicity by a study demonstrating resistance to acetaminophen-induced hepatotoxicity in GSTM1-null mice [21]. For understanding of the relation between GSTM3 and HD-MTX-induced hepatotoxicity, further research into the role of GSTM3 in the MTX processing pathway is required. The AGG deletion is reported to generate a recognition site for the transcription factor YY1, which is known to regulate gene expression from intragenic sites [22]. YY1 has been suggested to act as a GSTM3 inducer [23]. We therefore speculate that carrying the rs1799735 variant leads to increased glutathione levels, which by scavenging of radicals is protective for hepatotoxicity.

Of the genes found in association with MTX plasma levels, only ATP-Binding Cassette Sub-Family C (CFTR/MRP) Member 4 (ABCC4) is an established player in the metabolic route of MTX. ABCC4 is involved in the transport of various organic anions and exogenous agents (including MTX) out of the cell. Both ABCC2 and ABCC4 are involved in MTX clearance via the kidneys [24]. In addition, ABCC4 is expressed in hepatocytes, where it transports most of the MTX from the liver cells into the blood circulation. In line with our results, studies in ALL patients have already indicated associations of ABCC4 variants with MTX plasma levels, though both SNPs identified in the present study were not investigated in the previous studies [25,26]. The previously reported ABCC4 variants in association with MTX levels (rs9516519 and rs868853) were not investigated in the present study and are not in LD with any of the ABCC4 variants covered by the DMET array.

The top three of associated genes with 48 hour MTX plasma levels were all biotransformation enzymes: Microsomal Epoxide Hydrolase 1 (EPHX1), cytochrome P450 2B6 (CYP2B6) and cytochrome P450 2C8 (CYP2C8). EPHX1 is involved in the activation and detoxification of epoxides and is expressed in most tissues [27]. Two different variants in this gene, with opposing effects and not in LD, were significantly associated with the MTX plasma levels. The identified variants are not in LD with two other well-studied EPHX1 variants (rs1051740 and rs2235922) with a functional effect on the enzyme activity, of which the rs1051740 variant was present on the DMET array but was not related to MTX plasma levels in our study [28]. CYP2B6 and CYP2C8 both encode hepatic enzymes involved in the metabolism for a wide variety of drugs including antineoplastic agents, but both have not been previously linked to MTX [29,30].

The DMET Plus array includes the functional polymorphism in SLC01B1 (rs4149056), which is shown to be significantly associated with MTX clearance in a GWAS in patients with ALL and in a small osteosarcoma study [6-8]. Analyzing this variant in relation to MTX plasma levels in our osteosarcoma patients did not reveal this association (P=0.51).

This difference in studies might be explained by the different cohorts used (patients with ALL and osteosarcoma, respectively). Furthermore, different endpoints are used, being MTX clearance, mucositis and infection in the GWAS, and MTX plasma levels and increased liver enzymes in the present study. The osteosarcoma study by Goricar et al. has analyzed tumor DNA, which may also introduce differences in study results.

The depiction of hepatotoxicity in this study was based on the CTCAE criteria, which is commonly used in oncology. By considering only grade 3 and 4 as toxicity, we have chosen a more strict definition of hepatotoxicity (>5x upper limit of normal) than the Drug-Induced Liver Injury (DILI) guidelines (>3x upper limit of normal), to improve clinical relevance of the results in the context of treatment for which high levels of toxicity are accepted to improve patient survival [31].

The power of this exploratory study was limited because of the relatively low patient number for an association study. Therefore, we have performed a two-step approach, by first selecting genetic variants predictive of MTX plasma levels as an intermediary endpoint, and subsequently analyzing only significantly associated genetic variants for their relation with hepatotoxicity. Inclusion of genetic variants in the association analysis with hepatotoxicity was further restricted by selecting only genetic variants that were associated with both continuous and dichotomized data for MTX levels. Furthermore, to make optimal use of the available data, we have included all available repeated measurements of the MTX plasma levels in the analyses.

In conclusion, we have identified variants in genes associated with elevated MTX plasma levels and hepatotoxicity, which are largely unknown to play a role in these processes. Pinpointing the role of these genes in MTX processing will be important to understand interpatient variability in the MTX response and will subsequently lead to optimized MTX treatment for osteosarcoma patients. Replication and functional validation of the results in larger cohorts is needed, in order to make these results clinically applicable.
References


Pharmacogenetics of methotrexate hepatotoxicity


Drug Efficacy
Chapter 5

Pharmacogenetic risk score is predictive of treatment response in patients with osteosarcoma


* These authors contributed equally
Abstract

Despite multi-agent chemotherapeutical treatment, osteosarcoma patients relapse frequently while survival rates have not improved significantly over the past decades. We have previously identified five genetic markers that could distinguish patients with good outcome from those with poor outcome, which can potentially be used for risk stratification at diagnosis. Here, we assessed the reproducibility of these findings in three independent osteosarcoma cohorts.

A Dutch cohort (N=191), Spanish cohort (N=95), and English cohort (N=51) of patients with high-grade osteosarcoma were genotyped for FasL rs763110, MSH2 rs4638843, CASP3 rs2720376, ABCC5 rs939338, and CYP3A4 rs4646437. Differences in 5-year progression-free survival (PFS) based on the combined effect of these variants, calculated as a genetic risk score, were assessed by the log rank test. Associations of the individual variants with 5-year PFS were determined in a meta-analysis using Cox proportional hazards models, with a total of 446 patients (including data from our previous study, N=177), excluding the English cohort for which significantly superior PFS was observed.

Genotyping of all variants except for rs4646437 was successful in all cohorts. In the Dutch cohort, the genetic risk score showed marginally significant differences in 5-year PFS (P=0.062). Although not significant in the other cohorts separately, in the combined Dutch and Spanish cohort PFS significantly differed based on the number of risk alleles (P=0.040). Meta-analysis (of the Dutch and Spanish cohorts) also including the two original cohorts, confirmed significant associations of the individual variants in MSH2 and ABCC5 with 5-year PFS.

We confirmed the association of our previously reported genetic risk score with PFS in two independent cohorts of patients with osteosarcoma. This provides a rationale for prospective testing in larger cohorts and may eventually contribute to prediction of treatment outcome and optimization of therapy for patients with osteosarcoma.

Introduction

Osteosarcoma is a bone tumor that mainly affects children, adolescents, and young adults. Although rare, it is the most common primary bone sarcoma. The backbone of chemotherapy for osteosarcoma consists of cisplatin and doxorubicin and in some protocols methotrexate (MTX). The introduction of this multi-agent chemotherapy in the 1970s, in addition to surgery, has given a major improvement in survival rates from below 20% to 60-65% [1]. However, osteosarcoma patients still relapse frequently and no further improvements in survival have been achieved since [2]. The absence of reliable predictive factors for risk stratification has resulted in uniform treatment schedules for osteosarcoma patients [3]. Hence, there is a need to optimize current treatment strategies and to develop novel approaches for the treatment of osteosarcoma patients.

Insights into the pharmacogenetics of treatment response in patients with osteosarcoma may contribute to upfront risk stratification. In an exploratory study focusing on genes of the metabolic pathways of cisplatin and doxorubicin [4], we previously identified significant associations of five genetic markers (Fas Ligand (Fasl) rs763110, MutS homologue 2 (MSH2) rs4638843, Caspase 3 (CASP3) rs2720376, ATP-binding cassette sub-family C member 5 (ABCC5) rs939338, and Cytochrome P450 3A4 (CYP3A4) rs4646437) with treatment outcome. Combining these variants in a genetic risk score (GRS) was predictive of 5-year progression-free survival (PFS) in two combined cohorts of a total of 172 patients with osteosarcoma [4]. Further analysis using independent cohorts is important to confirm the associations and assess the robustness of these findings. Therefore, in the present study we assessed the reproducibility of the association of the GRS with 5-year PFS in three independent osteosarcoma cohorts from The Netherlands, Spain, and England.

Materials and methods

Patient cohorts

A Dutch cohort of 191 patients with osteosarcoma, unrelated to the cohorts investigated in the original study, was retrospectively recruited at the Radboud university medical center (Nijmegen), the University Medical Center of Groningen (Groningen), Leiden University Medical Center (Leiden) and the Emma Children’s Hospital/Academic Medical Center (Amsterdam), The Netherlands. All patients were treated between 1982 and 2013 and the clinical data was retrospectively collected from medical records. Eligibility required the following criteria: diagnosis of primary high-grade osteosarcoma with or without metastatic disease, age ≤45 years, treatment with cisplatin and doxorubicin-based chemotherapy, and self-reported Caucasian ethnicity. Patients were treated as
described previously [4]. The study was approved by the Institutional review board of the Radboud university medical center (Commissie Mensgebonden Onderzoek Regio Arnhem Nijmegen), and approval for inclusion of patients in other Dutch institutes was obtained from institutional ethics committees. All patients and/or parents provided written informed consent.

For confirmation of the positive findings in the Dutch cohort, two cohorts of high-grade osteosarcoma patients treated with cisplatin and doxorubicin-based chemotherapy from Spain (N=95, diagnosed between 1985 and 2014) and England (N=51, diagnosed between 2002 and 2008) were used. Information on treatment has been reported previously [5,6]. The same inclusion criteria as used in the Dutch cohort were applied, with the exception of ethnicity. The study was approved by the local ethics committees and written informed consent was obtained from parents and/or patients.

Genotyping methods
From the Dutch cohort, germline DNA was isolated from blood or saliva of patients alive, or from normal formalin-fixed paraffin-embedded (FFPE) tissue of deceased patients, as described previously [4]. Details on DNA isolation of the Spanish and English cohorts has been previously reported [5,6]. Genotyping of the cohorts was performed for five single nucleotide polymorphisms (SNPs) selected on the basis of their significant association with 5-year PFS in the meta-analysis of the two cohorts from the original study, and improved significance compared to the discovery cohort alone [4]. KASP-On-Demand (KOD) assays were used for FASL rs7631110 and CYP3A4 rs4646437; KASP-By-Design (KBD) assays were used for MSH2 rs4638843 and CASP3 rs2720376, all according to the manufacturer’s protocol (LGC Genomics, Hoddesdon, UK). The ABCC5 rs939338 variant was determined using a TaqMan allelic discrimination assay (C__8759155_20) according to the manufacturer’s protocol (ThermoFisher, Nieuwegein, The Netherlands). Fluorescence was measured with a 7500FAST Real-Time PCR System (ThermoFisher).

Genotypes were scored using the algorithm and software (v2.0.6) supplied by ThermoFisher. Blanks (3%) as well as duplicates between plates were included as quality controls for genotyping.

Statistical analysis
Statistical differences in demographic data between cohorts were assessed by the Fisher exact or Pearson chi-square test (in case of dichotomized variables), Kruskal-Wallis or Mann-Whitney U test (in case of linear variables), or Cox proportional hazards analysis (in case of PFS). Five-year PFS was defined as the interval between diagnosis and disease progression/recurrence (=event) and was estimated using the Kaplan–Meier method. Patients without disease recurrence at the date of last follow-up were censored at that date. Associations of clinical variables with 5-year PFS were assessed by Cox regression analysis. The data analysis plan involved two stages: (1) assessment of the reproducibility of the association of GRS with 5-year PFS in a cohort that is very similar to the original study cohorts (i.e. the Dutch cohort), and (2) further testing of the reproducibility of the GRS association in two individual cohorts from abroad (Spanish and English cohorts). The GRS was constructed as in our previous study; differences in 5-year PFS between the patients, grouped by their number of risk alleles, were assessed using the log rank test. Only patients successfully genotyped for all variants included in the GRS analysis were included in the GRS calculation. The predictive accuracy of the GRS was quantified using Harrell’s C-statistic (which can range from 0.5, indicating prediction equal to chance, to 1.0, indicating perfect prediction, with 0.8 and higher considered strong predictive accuracy), based on Cox regression analysis of the 5-year PFS with the GRS as covariable. Association analysis of the individual SNPs with 5-year PFS was carried out by multivariable Cox regression analysis including MTX treatment as covariable, in line with the original study. The results of the separate cohorts were meta-analyzed using a fixed effects or random effects (in case of large heterogeneity: I²>50) model in METAL [7]. All other analyses were performed using SPSS package version 22 (SPSS Inc., Chicago, Ill, USA). Reported P-values are two-sided and are considered statistically significant if <0.05.

Results
Patient population
The study was carried out in two stages, with as a first stage assessment of reproducibility of previous findings using a Dutch cohort which is similar to the original study cohorts, and subsequent confirmation in two different patient populations using a Spanish cohort and an English cohort. The patient demographics for each cohort are provided in Table 1. From the Dutch cohort, 17 of 191 patients were excluded based on failed genotyping for all variants, which might be explained by potentially poor quality FFPE tissue-derived DNA of 15 of these patients. Several differences were observed in baseline characteristics between the cohorts. The patients in the Spanish cohort were significantly younger than the patients in the Dutch and English cohorts (P<0.001). In addition, more patients were treated with high-dose MTX in the Spanish and English cohorts than in the Dutch cohort (P<0.001). Moreover, in the English cohort a poor histologic response was much less frequently observed than in the Dutch and Spanish cohorts (P<0.001), and 5-year PFS was ~20% higher compared to the other cohorts (P=0.037). This might be a reflection of differences in inclusion of patients, because in the English cohort the patients who had died (and most likely with poor PFS) before the time of retrospective patient recruitment were not included. For this reason, this is considered a bias in the selected population.
Only in the Dutch cohort, the poor histologic responders showed significantly inferior 5-year PFS ($P=0.001$); in all cohorts the presence of primary metastases was significantly associated with inferior 5-year PFS. Therefore, the GRS analysis also included stratification for metastases at diagnosis, in line with the original study. However, as the histologic response and PFS are both a reflection of the response to chemotherapy, the histologic response was not adjusted for as this would unintentionally remove variation resulting in overcorrection. Although not significantly associated with PFS in the different cohorts, treatment with MTX was included as covariable in the single SNP analysis, in line with the original study.

**Table 1.** Demographic data of patients included in the cohorts.

<table>
<thead>
<tr>
<th></th>
<th>Dutch cohort</th>
<th>Spanish cohort</th>
<th>English cohort</th>
<th>P-value overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis, median (range)</td>
<td>16.4 (3.4-45.8)</td>
<td>14.0 (9.9-34.4)</td>
<td>17.0 (10.0-39.8)</td>
<td>&lt;0.001$^a$</td>
</tr>
<tr>
<td>Male gender, n (%)</td>
<td>100 (57.5%)</td>
<td>53 (55.8%)</td>
<td>29 (56.9%)</td>
<td>0.965</td>
</tr>
<tr>
<td>Axial tumor, n (%)</td>
<td>7 (4.0%)</td>
<td>2 (2.1%)</td>
<td>3 (5.9%)</td>
<td>0.499</td>
</tr>
<tr>
<td>Primary metastases, n (%)</td>
<td>24 (13.8%)</td>
<td>17 (17.9%)</td>
<td>6 (11.8%)</td>
<td>0.538</td>
</tr>
<tr>
<td>MTX treatment$^b$, n (%)</td>
<td>90 (51.7%)</td>
<td>84 (95.5%)</td>
<td>51 (100%)</td>
<td>&lt;0.001$^a$</td>
</tr>
<tr>
<td>Poor histologic response$^c$, n (%)</td>
<td>84 (53.8%)</td>
<td>49 (55.1%)</td>
<td>11 (22.0%)</td>
<td>&lt;0.001$^c$</td>
</tr>
<tr>
<td>5-year PFS</td>
<td>55.5%</td>
<td>58.2%</td>
<td>74.5%</td>
<td>0.007$^c$</td>
</tr>
</tbody>
</table>

MTX, methotrexate; PFS, progression-free survival.

$^a$ Data available on MTX treatment for each cohort: Dutch N=174, Spanish N=88, English N=51.

$^b$ Data available on histologic response for each cohort: Dutch N=156, Spanish N=89, English N=50.

$^c$ Ethnicities Spanish cohort: Caucasian N=90 (94.7%), Latin American N=4 (4.2%), Romani N=1 (1.1%).

$^d$ Ethnicities English cohort: Caucasian N=36 (70.6%), Afro-Caribbean N=8 (15.7%), Indian/Asian N=7 (13.3%).

$^e$ P-values are given for overall tests including all three cohorts. Significant associations ($P<0.05$): 

Genotypic

Of the five variants included in the genotyping, the CYP3A4 rs4646437 variant was not reliably determined, most likely because of close homology of the primer binding region with another genomic region. The average call rate of the four remaining variants was 0.95. FASL rs763110 and ABCG5 rs939338 were not in Hardy Weinberg equilibrium ($P=0.032$ for both variants), in both cases driven by the Dutch cohort. However, this was not considered a reason for exclusion because this might be attributed to the number of drop-outs and/or to analysis in a diseased population [8]. Moreover, samples genotyped in duplicate showed positive results, and genotypes of 92 additional healthy control samples were in HWE for both variants. No major differences in minor allele frequencies were observed between the cohorts (Supplementary Table S1).

**Genetic risk score analysis**

The genetic risk score was calculated using the four variants rs763110, rs4638843, rs2720376, and rs939338 (for comparison with the PFS of the GRS of the four variants in the original study, see Supplementary Figure S1). After grouping of patients by the number of risk alleles similar to the original study (0-1, 2-3, 4-5, more than 5 risk alleles), we observed marginally significant differences in 5-year PFS in the Dutch cohort ($P=0.062$) (Figure 1A). Whereas patients carrying no or one risk allele showed a 5-year PFS of 72.4%, and 66.2% with 2-3 risk alleles, the patients carrying 4-5 risk alleles showed significantly inferior PFS of 42.1% ($P=0.042$ compared to 0-1 allele; $P=0.023$ compared to 2-3 alleles). Subsequent analysis in the Spanish validation cohort did not reveal overall significant differences in PFS, although the survival curves of the different GRS groups also showed divergence between the groups with the smallest and largest number of risk alleles (Figure 1B). Though not significant, the opposite effect of the GRS on PFS was observed in the English cohort (Supplementary Figure S2), which is also reflected in the combined analysis of all three cohorts, in which the PFS differences between the GRS groups were absent (Figure 1C). The observation of the opposite effect in the English cohort may be attributed to the aforementioned differences in PFS and selection of patients that were included between the English cohort and the Dutch and Spanish cohort. After analysis of the data, we recognized that the English cohort was not appropriate for validation because of the combination of its selected population and small sample size; however, we chose to follow the original analysis plan for GRS analysis in the individual cohorts. Furthermore, because of the discrepancy in PFS, we carried out a combined analysis including only the Dutch and Spanish cohort, which revealed significant differences in 5-year PFS between the GRS groups ($P=0.040$) (Figure 1D). Patients carrying no or one risk allele showed a 5-year PFS of 78.5%, compared to a 5-year PFS of 59.5% in patients carrying 2-3 risk alleles, 48.0% with 4-5 alleles, and 57.1% with more than 5 risk alleles ($P=0.040$). After stratification based on metastatic status at diagnosis, the patients without metastases at diagnosis showed a significant difference in PFS based on the GRS ($P=0.021$) with a pattern similar to the complete group of patients (Supplementary Figure S3). In the patients with metastases at diagnosis, the GRS was not significantly associated with PFS, although markedly inferior PFS was observed in each GRS group compared to the complete group of patients, with the exception of the one patient carrying more than 5 risk alleles.

The predictive accuracy of the GRS was determined combining the two cohorts from the original study with the Dutch and Spanish cohorts of the present study, resulting in a C-statistic of 0.61.
Confirmation of pharmacogenetic risk score in osteosarcoma

Figure 1. Five-year PFS based on GRS. A: Dutch cohort (N=150), patients with 0-1 risk allele showed a 5-year PFS of 72.4% compared to 66.2% with 2–3 risk alleles, 42.1% with 4–5 alleles, and 75.0% with >5 risk alleles. B: Spanish cohort (N=90), patients with 0-1 risk allele showed a 5-year PFS of 90.0% compared to 50.3% with 2–3 risk alleles, 60.0% with 4–5 alleles, and 33.3% with >5 risk alleles. C: Dutch, Spanish, and English cohort combined (N=283), patients with 0-1 risk allele showed a 5-year PFS of 71.1% compared to 63.1% with 2–3 risk alleles, 55.9% with 4–5 alleles, and 62.5% with >5 risk alleles. D: Dutch and Spanish cohort combined (N=240), patients with 0-1 risk allele showed a 5-year PFS of 78.5% compared to 59.5% with 2–3 risk alleles, 48.0% with 4–5 alleles, and 57.1% with >5 risk alleles; with significant difference between the 0-1 and 4-5 risk allele groups (P=0.006).

Meta-analysis associations of individual SNPs

Association analysis of the individual SNPs with 5-year PFS was carried out using multivariable Cox regression analysis, including MTX treatment as covariable, similar to the original study. In meta-analysis of the results of the two cohorts of the original study (N=177) and the Dutch and Spanish cohorts, resulting in a total population of 446 patients, significant associations with 5-year PFS were observed for MSH2 rs4638843 (HR 1.63 [1.21-2.20]; P=0.002) and ABCC5 rs939338 (HR 1.36 [1.10-1.68]; P=0.005), with the same direction of effect in all cohorts and increased significance compared to the results of the meta-analysis of the original cohorts only (Table 2).

Table 2. Associations with 5-year PFS in combined meta-analysis of original study cohorts and Dutch and Spanish cohorts.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene</th>
<th>Allele</th>
<th>HR</th>
<th>95% CI</th>
<th>P-value</th>
<th>I²</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs4638843</td>
<td>MSH2</td>
<td>C</td>
<td>1.63</td>
<td>1.21-2.20</td>
<td>0.002</td>
<td>5</td>
</tr>
<tr>
<td>rs939338</td>
<td>ABCC5</td>
<td>G</td>
<td>1.36</td>
<td>1.10-1.68</td>
<td>0.005</td>
<td>0</td>
</tr>
<tr>
<td>rs263110</td>
<td>FASL</td>
<td>T</td>
<td>1.23</td>
<td>0.99-1.53</td>
<td>0.065</td>
<td>20.7</td>
</tr>
<tr>
<td>rs2720376</td>
<td>CASP3</td>
<td>G</td>
<td>0.80</td>
<td>0.55-1.15</td>
<td>0.22</td>
<td>59.4</td>
</tr>
</tbody>
</table>

95% CI, 95% confidence interval; HR, hazard ratio. HR and 95% CI are reported for the minor allele, a HR<1 indicates a protective effect of the minor allele and risk of an event for the major allele. Meta-analysis for each of the variants was performed with a fixed effects or random effects model depending on the corresponding heterogeneity (I²) as indicated in the table.

Discussion

One of the current challenges on the road to optimization of treatment of osteosarcoma is to identify those patients at risk of a poor outcome. We have previously explored several pharmacogenetic markers associated with treatment response in a cohort of osteosarcoma patients [4]. In the present study, we have confirmed the association of a genetic risk score based on these markers (FasL rs763110, MSH2 rs4638843, CASP3 rs2720376, ABCC5 rs939338) with 5-year PFS in independent cohorts. In addition, we have confirmed the association of two of the individual SNPs with PFS in a meta-analysis also including the two cohorts of the original study.

In the current study, we were unable to analyze the role of the rs4646437 marker located in the CYP3A4 gene, which encodes the CYP3A4 enzyme involved in the metabolism of both cisplatin and doxorubicin. However, even with exclusion of this marker from the genetic risk score calculation, the association of the score with 5-year PFS remained
significant in the patient group of the original study, and the same trend was observed in the Dutch cohort, reaching significance in the combined Dutch and Spanish cohort.

Both the GRS analysis and meta-analysis were carried out with the (combined) Dutch and Spanish cohorts, excluding the English cohort. The reason for this was the discrepancy of the PFS in the English cohort compared to the other cohorts, probably due to its relatively small sample size (51 eligible patients, 43 patients successfully genotyped for the four genetic variants and therefore included in the GRS analysis) and approach in patient inclusion, i.e. patients who had died before the start of retrospective patient recruitment were not included in the study. Unfortunately, this may indicate a selection of patients with relatively favorable PFS. Probably, inclusion of the deceased patients, which would result in more events, would be needed to obtain conclusive results for this cohort. In addition to the expected overrepresentation of survivors, the English cohort is also different from the other cohorts with respect to time frame of inclusion, MTX treatment, and ethnic background. However, increases in survival due to for example improved supportive care in the more recent time frame are often relatively subtle, and the proportion of patients with overall MTX treatment is similar to the Spanish cohort. Furthermore, (overall) survival differences between ethnicities, if any, are directed towards inferior survival in black patients [9]. Therefore, these factors are not expected to have substantial impact on the 5-year PFS such as observed in the English cohort.

Most pharmacogenetic studies in osteosarcoma to date have not been reproduced or have shown inconsistent findings. In general, demonstration of reproducibility of pharmacogenetic findings is of vital importance in the process to clinical implementation, but often not achieved [10]. The approach of the current study was to assess the reproducibility of our previous results in a cohort similar to the cohorts in the original study with regard to the approach and criteria of inclusion (also of deceased patients), inclusion sites, treatment protocols, and ethnicity (i.e. only Caucasian origin), as a first step to assess the relevance of the findings. This was extended to assessment of reproducibility of the findings in cohorts that differ from the original study, with regard to inclusion sites, treatment protocols, and ethnicity, as a separate stage. Following this approach, we have been able to validate our previously reported finding of the association of the GRS with 5-year PFS. This was not only the case in the Dutch cohort, but also to some extent in the Spanish cohort, indicating its robustness. Although the association of the GRS with PFS was not statistically significant in the Spanish cohort alone, the differences in PFS between the risk score groups were in similar direction as in the Dutch cohort, which is also reflected by the significant association after combining both cohorts. Although these results are promising, the reproducibility was not complete for all risk score groups as the PFS of the small proportion of patients carrying more than 5 risk alleles was not significantly different in the analyses and was even better in the Dutch cohort compared to patients carrying 4 to 5 risk alleles. This might be related to the small sample size of the group of patients carrying more than 5 risk alleles and/or the fact that these patients also carried at least two risk alleles of the two variants not individually confirmed in the meta-analysis, which are therefore not expected to have a dominant contribution to the PFS. Furthermore, the association of the GRS with PFS could not be confirmed in the English cohort, indicating that the association may not be 100% robust, although it should be noted that significant differences in PFS were observed as indicated above.

As patients with localized or metastatic disease are known to have a different prognosis, also in the three cohorts in this study, stratification on the basis of metastatic status at diagnosis was performed. The results indicated that this improved the ability to differentiate patients based on the GRS; with overall inferior PFS for each risk score group in metastatic patients compared to patients without metastases at diagnosis. However, also here the analysis of patients carrying more than 5 risk alleles was limited because of low numbers, and this should be further investigated in larger populations.

The association of two of the individual variants (MSH2 rs4638843 and ABCC5 rs939338) with 5-year PFS was validated in a meta-analysis. Therefore, these may be especially interesting candidates to follow-up for functional characterization. Both genes were selected for analysis in the original pathway-based study based on the described relation to cisplatin and/or doxorubicin metabolism and transport [4]. Biological validation could be useful to determine whether the identified variants are the causal variants themselves or whether they are linked to the actual causal variant. In addition, it could provide additional evidence and a biological rationale for further research towards translation to a clinical setting. Furthermore, it may give more insight into the mechanisms underlying the drug response.

Although the number of patients included in this study is relatively high for osteosarcoma because of its rarity, the number of patients is rather low for validation studies. This may have limited the ability to significantly confirm the associations in each individual cohort. Therefore, we suggest further confirmatory studies in larger (prospective) cohorts, before clinical implementation of the risk score. In addition, the predictive accuracy of the genetic risk score based on the four identified variants is relatively low, indicating that these specific pharmacogenetic variants alone are not sufficient to fully discriminate patients prone to have a poor treatment response from those with a good outcome. This implicates that additional (non) genetic risk factors would be needed to be able to construct a predictive model. The selection of genes for the original study was based on the known role in metabolic routes of cisplatin or doxorubicin. On the basis of the current consistent results, indicating that indeed pharmacogenetics may play a role in the treatment response in osteosarcoma, further research into genes not directly related to these agents would be worthwhile. As it is most likely that the tumor genetics and biology contribute to variation in treatment response, this should also be further characterized [11]. In the future, the findings of the current study, combined with additional predictive
factors, may provide a clinical tool for risk stratification. The next step will be to consider what the alternative options would be for patients prone to have a poor response to standard chemotherapy.

In conclusion, we have confirmed the association of a genetic risk score based on four pharmacogenetic variants with PFS in patients with osteosarcoma. This may provide a basis for pharmacogenetic profiling in the future and could contribute to treatment optimization for patients with osteosarcoma.

References

Confirmation of pharmacogenetic risk score in osteosarcoma

Supplementary data

Figure S1. Five-year PFS of the original study cohort (N=172) based on GRS calculated without the CYP3A4 rs4646437 variant. Patients with 0-1 risk allele showed a 5-year PFS of 80.2% compared to 69.4% with 2-3 risk alleles, 46.4% with 4-5 alleles, and 16.7% with >5 risk alleles.

Figure S2. Five-year PFS based on GRS in the English cohort (N=43). Patients with 0-1 risk allele showed a 5-year PFS of 42.9% compared to 73.7% with 2-3 risk alleles, 87.5% with 4-5 alleles, and 100% with >5 risk alleles.

Figure S3. Five-year PFS based on GRS of Dutch and Spanish cohort combined, stratified for metastatic status at diagnosis. A: Patients without metastases and with 0-1 risk allele showed a 5-year PFS of 82.9% compared to 65.3% with 2-3 risk alleles, 49.3% with 4-5 alleles, and 50.0% with >5 risk alleles. B: Patients with metastases and with 0-1 risk allele showed a 5-year PFS of 53.3% compared to 21.1% with 2-3 risk alleles, 41.7% with 4-5 alleles, and 100% with >5 risk alleles.

Table S1. Minor allele frequencies of the four variants in each cohort.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Allele</th>
<th>Dutch cohort (N=174)</th>
<th>Spanish cohort (N=95)</th>
<th>English cohort (N=51)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs762110</td>
<td>T</td>
<td>0.338</td>
<td>0.378</td>
<td>0.439</td>
</tr>
<tr>
<td>rs638845</td>
<td>C</td>
<td>0.111</td>
<td>0.139</td>
<td>0.089</td>
</tr>
<tr>
<td>rs939338</td>
<td>G</td>
<td>0.409</td>
<td>0.332</td>
<td>0.480</td>
</tr>
<tr>
<td>rs2780376</td>
<td>G</td>
<td>0.416</td>
<td>0.457</td>
<td>0.480</td>
</tr>
</tbody>
</table>
Chapter 6

Genetic variants of drug transporters influence response to chemotherapy-based treatment in osteosarcoma


* These authors contributed equally

Submitted
Abstract

Upfront prediction of efficacy of chemotherapy for osteosarcoma may enable treatment optimization and improvement of outcome. Pharmacogenetic studies of the cisplatin and doxorubicin metabolic pathways have indicated a role of genetic variants in treatment response in osteosarcoma. As the complex metabolism of drugs used in osteosarcoma treatment may involve a broader range of genetic factors, we investigated a comprehensive set of 1,936 variants in 231 genes known to be involved in drug metabolism and transport.

In a discovery stage including two Dutch cohorts of high-grade osteosarcoma patients (N=139 and N=177), germline DNA was genotyped using the Drug Metabolizing Enzymes and Transporters (DMET) Plus array. Associations of genetic variants with histologic response to preoperative cisplatin and doxorubicin-based chemotherapy, and with 5-year progression-free survival (PFS) were evaluated using multivariable logistic regression models and Cox proportional hazards models, respectively. Statistically significant findings were validated in Spanish (N=100) and English (N=54) cohorts.

After meta-analysis of the Dutch cohorts, significant associations of 23 genetic variants with 5-year PFS were identified. Of these markers, the ABCB1 rs17064 variant was independently validated in the Spanish cohort (P=0.014), and showed increased significance in association with 5-year PFS in a combined meta-analysis of the Dutch and Spanish cohorts (HR for T allele 1.89, P=0.001), excluding the English cohort which showed markedly higher PFS. Seven variants in five other genes (ABCC4, ABCC5, ABCG1, SLC22A5, and SLC22A14) were also significantly associated with PFS in combined meta-analysis of the Dutch and Spanish cohorts. With addition of the English cohort to the combined meta-analysis, six of the eight variants remained significantly associated with PFS. Of 15 genetic variants associated with histologic response in the discovery phase, the top associated variants could not be confirmed in the validation cohorts.

This is the first pharmacogenetic study in osteosarcoma patients employing a comprehensive absorption, distribution, metabolism, and excretion (ADME) pathway analysis in relation to PFS. We have identified associations of genetic variants of multiple drug transporters with outcome after chemotherapy-based treatment. After validation in larger cohorts, functional validation and prospective testing, these markers are of potential interest for the development of new treatment strategies and optimizing current therapy.

Introduction

In patients with osteosarcoma, no substantial improvements in outcome after chemotherapy-based treatment have been achieved in the past decades. Treatment consists of a combination of surgery and chemotherapy, mostly based on cisplatin, doxorubicin, and in some protocols methotrexate (MTX), and is similar for all patients. Although this combination of drugs is the most effective one and used in various protocols, the treatment is successful in a subset of patients, with cure rates approaching 60-70% for patients with non-metastatic disease and 20-40% for patients with metastatic disease [1]. To improve outcome, upfront identification of patients at risk of a poor therapy response is essential to further optimize therapy.

Metastatic status at diagnosis is the only consistent prognostic factor, but is not sufficient to explain the complete range of differences in outcome. The histologic response after preoperative chemotherapy is often used as a measure of drug response at the time of surgery, but it is not available for risk stratification at diagnosis and has not been confirmed as a consistent predictor of treatment outcome [2,3]. As interindividual variation in drug response is in part determined by genetic variation, pharmacogenetic approaches to identify genetic predictors of treatment response in osteosarcoma have gained attention in recent years. Multiple candidate gene studies as well as a few pathway-based gene studies have suggested a role of pharmacogenetic markers in the treatment response in osteosarcoma [4,5]. In a previous pathway-based gene study from our group, covering the metabolic pathways of cisplatin and doxorubicin, we identified candidate genes associated with 5-year progression-free survival (PFS) [6]. Nevertheless, the presence of these genetic variants was not sufficient to fully predict treatment response, suggesting the involvement of additional genes. Therefore, we set out to assess a more extensive set of genes involved in drug metabolism and transport, including genes not previously connected to cisplatin or doxorubicin metabolism. We performed a large scale screening of 1,936 genetic variants in 231 drug metabolism and transporter genes in Dutch osteosarcoma patients to investigate associations with treatment response, followed by validation in independent cohorts of osteosarcoma patients.

Materials and methods

Patient cohorts

For discovery, two Dutch cohorts of osteosarcoma patients were included: cohort I with 139 patients retrospectively recruited at the Radboud university medical center (Nijmegen) and the University Medical Center of Groningen (Groningen); cohort II with
177 patients from Leiden University Medical Center (Leiden) and the Emma Children’s Hospital/Academic Medical Center (Amsterdam), The Netherlands. These cohorts were based on the two cohorts investigated in our previous pathway-based study, with addition of an inclusion site (Amsterdam) to cohort II and additional patients because of ongoing recruitment at the other sites [6]. All patients were diagnosed between 1975 and 2013; clinical data was retrospectively collected from medical records. Patients were included if they were diagnosed with primary high-grade osteosarcoma with or without metastatic disease, aged ≤45 years, treated with cisplatin and doxorubicin-based chemotherapy, and of self-reported Caucasian ethnicity. Treatment of patients was either according to institutional standard therapy consisting of cisplatin (maximum cumulative dose 600 mg/m²) and doxorubicin (maximum cumulative dose 450 mg/m²), or according to the standard schedule as given in the EURAMOS-1 trial, which consisted of cisplatin (maximum cumulative dose 480 mg/m²), doxorubicin (maximum cumulative dose 450 mg/m²) and additionally high-dose MTX (maximum cumulative dose 144 g/m²), with or without additional ifosfamide and etoposide, or interferon-α [7]. The study was approved by the institutional review board of the Radboud university medical center, and approval for inclusion of patients in other Dutch institutions was obtained from institutional ethics committees. Written informed consent was provided by all patients and/or parents.

Two high-grade osteosarcoma cohorts treated with cisplatin and doxorubicin-based chemotherapy from Spain (N=100, diagnosed between 1985 and 2014) and England (N=54, diagnosed between 2002 and 2008) were used for validation of the significant findings in the Dutch cohorts. Information on treatment has been reported previously [4,5]. The same inclusion criteria were applied as used in the Dutch cohorts, except for ethnicity. The study was approved by the local ethics committees and written informed consent was provided by patients and/or parents.

**Genotyping methods**

From the two Dutch cohorts, germline DNA was isolated from blood or saliva of patients alive, or from formalin-fixed paraffin-embedded (FFPE) tissue of deceased patients, as described previously [8]. The DNA samples were genotyped for 1,936 genetic variants using the Affymetrix DMET Plus array according to the manufacturer’s instructions (Affymetrix UK Ltd, High Wycombe, UK). Genotypes were calculated with DMET console software 1.3 using the Dynamic Genotype Boundaries version 2 algorithm. Variants were excluded from association analysis if the genotype cluster plots were considered unreliable, being plots with genotype calling showing merged clusters without distinct cluster boundaries. Additional stringent evaluation of the genotype clustering in combination with expected genotype frequencies was carried out for variants significant in association analysis. Samples and variants were excluded from analysis if call rates <0.9, minor allele frequency (MAF) <0.01 and/or deviating from Hardy-Weinberg equilibrium (HWE) (P-value <0.0001). Five copy number variants, 46 X-chromosomal variants and one tri-allelic variant present on the array were not included in the analyses.

Information on germline DNA isolation of the English and Spanish validation cohorts has been previously reported [4,5]. Genotyping of the validation cohorts was carried out for 20 of the 23 variants that were significantly associated with 5-year PFS in the discovery stage (P<0.05), excluding three variants that were in linkage disequilibrium (on the basis of r²=0.8) with any of the 20 variants and that had higher P-values than their linked variants. In addition, the validation cohorts were genotyped for the top three genes significantly associated with histologic response (P<0.01) in the Dutch cohorts, including one variant (SLC25A27 rs9381468) that was genotyped for association analysis of both endpoints. A more stringent significance threshold was applied for selection of variants relevant for validation of association with histologic response, because this endpoint may be less robust as it may be more sensitive to interobserver and sampling differences, in spite of standard definitions. For the validation cohort the following genetic variants were genotyped with KASP-On-Demand (KOD) assays: CD5318202, CHST5 rs2641806, SLC22A14 rs171248, SLC22A5 rs274548, SLC22A5 rs274558, SLC22A7 rs9381468, and UGT2B15 rs1902023; KASP-By-Design (KBD) assays were used for ABCB1 rs17064, ABCB4 rs2109505, ABCC4 rs3742106, ABCCS rs562, ABCCS rs7636910, ABCC1 rs3788007, CASP rs1364182, CYP39A1 rs7761731, FMO3 rs2266782, SULT1E1 rs38221172, and VKORC1 rs2884737, all according to the manufacturer’s protocol (LGCGenomics, Hoddesdon, UK). The ABCB2 rs2273697 (C____2272980____20), ABCCC2 rs717620 (C____2814642____10), ABCCC5 rs939338 (C____8759155____20), and SLC7A8 rs2268877 (C____2486192____1) variants were determined using a TaqMan allelic discrimination assay according to the manufacturer’s protocol (ThermoFisher, Newgeine, The Netherlands). For validation of the results for ABCCC5 rs939336, the closely linked ABCCC5 rs939338 variant (r²=0.98) was determined, as this was one of the variants that was also associated with 5-year PFS in our previous study [6]. Fluorescence was measured with a 7500FAST Real-Time PCR System (ThermoFisher). Genotypes were scored using the algorithm and software (v2.0.6) supplied by ThermoFisher. Blanks (3%) as well as duplicates between plates were included as quality controls for genotyping.

**Statistical analysis**

Statistical differences in demographic data between cohorts were assessed by the Fisher exact, Pearson chi-square, Mann-Whitney U test, or Cox proportional hazards analysis as appropriate. Five-year PFS was defined as the interval between diagnosis and disease progression/recurrence (=event) and was estimated using the Kaplan–Meier method. Patients without disease recurrence at the date of last follow-up were censored at that date. For analysis of the histologic response, a good response was defined as
SNPs in ADME genes and osteosarcoma treatment response

≥90% necrosis after preoperative chemotherapy. Associations of clinical variables with 5-year PFS were assessed by Cox regression analysis; for histologic response the Fisher exact or Pearson chi-square test were used.

The data analysis plan involved association analysis of the individual SNPs in both Dutch cohorts, and subsequent validation in each of the validation cohorts. The discovery stage included both Dutch cohorts that were analyzed using the complete DMET data and subsequently meta-analyzed for identification of overall significant associations. After association analysis of the positive findings from the discovery stage in the individual validation cohorts, meta-analysis of both discovery and validation cohorts was subsequently carried out if any of the tested variants showed Pr<0.1 in the validation cohorts.

To assess the effect of a genetic variant on the histologic response, the data were dichotomized to good/poor response. Associations between genetic variants and the histologic response were assessed by multivariable logistic regression analysis in PLINK using the command --logistic (additive model) [9]. Associations between genetic variants and 5-year PFS were assessed using Cox proportional hazards models using the specialized GenABEL package (additive model) [10] in R (version 3.1.0, http://www.r-project.org), which we validated using SPSS v22 (SPSS Inc., Chicago, Ill, USA). The results of the separate cohorts were meta-analyzed using a fixed-effects or random-effects (in case of large heterogeneity: I2 >50) model in METAL [11].

A genetic risk score (GRS) was constructed as in our previous study [6] based on genetic variants significantly associated with 5-year PFS; differences in 5-year PFS between the patients, grouped by their number of risk alleles, were assessed using the log rank test. Only patients successfully genotyped for each variant were included in the GRS calculation. All analyses were performed using SPSS unless stated otherwise. Reported P-values are two-sided and are considered statistically significant if <0.05. No correction for multiple testing was applied because of the exploratory nature of the discovery stage.

Results

Patient population

The study was carried out following a two-stage approach, with a discovery stage using two Dutch cohorts, and independent validation in a Spanish cohort and an English cohort. From both Dutch cohorts, three patients were excluded based on a genotype call rate lower than 0.9, resulting in 136 and 174 patients in the Dutch cohorts I and II, respectively. The clinical characteristics of patients in each of the four cohorts are provided in Table 1. From Dutch cohort I, 18 of 136 patients were excluded from association analysis of 5-year PFS because of ongoing treatment (N=1) or because they were not treated with neoadjuvant cisplatin and doxorubicin (N=17); from Dutch cohort II, 9 of 174 patients were excluded for the latter reason. From the validation cohorts, in total 8 patients were excluded from association analysis of 5-year PFS because they were not treated with neoadjuvant chemotherapy or not specifically with neoadjuvant cisplatin and/or doxorubicin (Spanish cohort N=4, English cohort N=3) or because no follow-up data was available (Spanish cohort N=1). No substantial differences were observed in baseline characteristics between the total cohorts and the patients eligible for 5-year PFS analysis (for comparison see Supplementary Table S1).

The Dutch cohorts significantly differed with respect to age at diagnosis, MTX treatment, and histologic response (Table 1). In Dutch cohort I, patients were slightly younger, more often treated with (neoadjuvant) MTX (because of differences in standard treatment protocols) and a poor histologic response was less frequently observed compared to Dutch cohort II. Despite the difference in poor responders, the Dutch cohorts did not significantly differ with regard to 5-year PFS. Due to these baseline differences, the genetic analyses of the Dutch cohorts were carried out using a meta-analysis. Furthermore, the English cohort had a significantly lower percentage of poor histologic responders (22.0%) compared to the other cohorts, 5-year PFS was ~15% higher than observed in the Dutch and Spanish cohorts (Supplementary Figure S1), and 5-year overall survival was superior to the other cohorts. Especially the high 5-year PFS in the English cohort may be a reflection of differences in inclusion of patients, as in the English cohort patients who were deceased (and most likely showing poor PFS) at the time of retrospective patient recruitment were not included. This is therefore considered a bias in the selected population. With respect to the histologic response, all patients of the English cohort were treated with HD-MTX preoperatively, which might have an influence on the histologic response, as indicated below.

In both Dutch cohorts, no significant differences were observed between good and poor responders with regard to the number of neoadjuvant cycles of cisplatin and doxorubicin and/or neoadjuvant cisplatin and doxorubicin dosage. In addition, none of the clinical variables included in Table 1 were significantly associated with histologic response in all cohorts. However, neoadjuvant MTX treatment was predictive of a good histologic response in cohort I (P=0.021). Because of the potential influence of MTX treatment on the histologic response, which may also partially explain the above-mentioned difference in histologic response between Dutch cohort I and II, neoadjuvant treatment with MTX was included as clinical covariable in the genetic association analyses of histologic response. Analysis of the clinical variables with 5-year PFS revealed a significant association of the presence of metastases at diagnosis with inferior PFS (P=0.002 and P=0.001 for cohort I and II, respectively); the same association was observed in the validation cohorts. As no other clinical variables were significantly associated with PFS in
all cohorts, only metastatic status at diagnosis was included as clinical covariable in the genetic association analyses of PFS. The association of a poor histologic response with inferior PFS was only significant in Dutch cohort II (P=0.001) and marginally significant in Dutch cohort I (P=0.06). Adjusting for the histologic response would unintentionally remove variation resulting in overcorrection, as histologic response and PFS are both a measure for the treatment response.

### Table 1. Demographic data of all patients in the discovery and validation cohorts, N=464.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Dutch cohort I (N=136)</th>
<th>Dutch cohort II (N=174)</th>
<th>Spanish cohort (N=100)</th>
<th>English cohort (N=54)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis, median (range)</td>
<td>15.7 (5.6-41.1)</td>
<td>17.5 (3.4-45.8)</td>
<td>14.0 (1.9-34.6)</td>
<td>17.5 (10.0-44.0)</td>
</tr>
<tr>
<td>Male gender, n (%)</td>
<td>73 (53.7)</td>
<td>98 (56.3)</td>
<td>55 (55.0)</td>
<td>31 (57.4)</td>
</tr>
<tr>
<td>Axial tumor, n (%)</td>
<td>5 (3.7)</td>
<td>7 (4.0)</td>
<td>3 (3.0)</td>
<td>3 (5.6)</td>
</tr>
<tr>
<td>Primary metastases, n (%)</td>
<td>23 (16.9)</td>
<td>25 (14.4)</td>
<td>19 (19.0)</td>
<td>7 (13.0)</td>
</tr>
<tr>
<td>MTX total, n (%)</td>
<td>92 (67.6)</td>
<td>80 (46.0)</td>
<td>88 (84.4)</td>
<td>51 (100)</td>
</tr>
<tr>
<td>MTX neoadjuvant, n (%)</td>
<td>73 (53.7)</td>
<td>50 (40.2)</td>
<td>32 (36.0)</td>
<td>50 (100)</td>
</tr>
<tr>
<td>Poor histologic response, n (%)</td>
<td>81 (60.8)</td>
<td>75 (48.7)</td>
<td>40 (54.4)</td>
<td>11 (22.0)</td>
</tr>
<tr>
<td>5-year PFS (%)</td>
<td>60.5%</td>
<td>60.8%</td>
<td>58.2%</td>
<td>74.5%</td>
</tr>
<tr>
<td>5-year OS (%)</td>
<td>76.3%</td>
<td>75.9%</td>
<td>82.4%</td>
<td>88.9%</td>
</tr>
</tbody>
</table>

MTX, methotrexate; PFS, progression-free survival; OS, overall survival.

* Dutch cohorts I and II significantly differed with respect to age at diagnosis (P=0.038), MTX treatment total (P<0.001) and neoadjuvant (P=0.018), and poor histologic response (P=0.001).

** Ethnicities English cohort: Caucasian N=38 (70.4%), Afro-Caribbean N=8 (14.8%), Indian/Asian N=8 (14.8%); Spanish cohort: Caucasian N=94 (94.0%), Latin American N=5 (5.0%), Romani N=1 (1.0%).

* Data available on MTX treatment total (neoadjuvant and/or adjuvant): Dutch cohorts all patients, Spanish cohort N=88, English cohort N=51.

* Data available on MTX treatment neoadjuvant: Dutch cohorts all patients, Spanish cohort N=89, English cohort N=50.


**Association analysis of histologic response**

Of the 1,884 variants included in the quality control of both Dutch cohorts, 90 variants were considered unreliable after primary screening of clusterplots. In addition, 36 (cohort I) and 32 (cohort II) variants were excluded because of genotype call rate <0.9, and 1,069 (cohort I) and 1,093 (cohort II) variants because of MAF <0.01. All remaining variants were in HWE. After quality control, 689 and 669 variants were included in the association analyses of Dutch cohort I and cohort II, respectively. In multivariable analysis of genetic variants and histologic response, including neoadjuvant MTX treatment as covariable, 18 genetic variants in 12 genes showed significant association with histologic response in meta-analysis of the results of the two Dutch cohorts. After stringent evaluation of the genotype clusterplots, three of these variants (GSTO1 rs4925, and VKORC1 rs2359612 and rs9923231) were excluded, resulting in 15 remaining variants in 11 genes (Supplementary Table S2). All variants showed the same direction of effect in the two Dutch cohorts.

To validate the most significant findings of the meta-analysis of the Dutch cohorts, three variants were investigated in the Spanish and English validation cohorts. The three variants, ABC2 rs717620, SLC25A27 rs9381468, and VKORC1 rs2884737, showed an average call rate of 0.96 and were in HWE (P>0.05). The validation cohorts were analyzed similar to the Dutch cohorts, although in the English cohort no correction for MTX treatment was applied as all patients received MTX. None of the variants were validated in association with histologic response. The T allele of the ABC2 rs717620 variant was associated with poor histologic response in the English cohort as reported previously [4]; the opposite effect compared to the findings in the Dutch cohorts.

**Table 2. Association analysis results of top associated variants with histologic response in discovery and validation stage.**

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene</th>
<th>Chr</th>
<th>Minor allele</th>
<th>OR</th>
<th>95% CI</th>
<th>p</th>
<th>OR</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs717620</td>
<td>ABC2</td>
<td>10</td>
<td>T</td>
<td>1.23</td>
<td>1.00-1.50</td>
<td>0.04</td>
<td>1.29</td>
<td>1.03-1.66</td>
<td>0.04</td>
</tr>
<tr>
<td>rs9381468</td>
<td>SLC25A27</td>
<td>6</td>
<td>T</td>
<td>1.24</td>
<td>1.13-1.36</td>
<td>0.002</td>
<td>0.98</td>
<td>0.90-1.09</td>
<td>0.05</td>
</tr>
<tr>
<td>rs2884737</td>
<td>VKORC1</td>
<td>16</td>
<td>G</td>
<td>0.57</td>
<td>0.49-0.66</td>
<td>0.001</td>
<td>0.50</td>
<td>0.36-0.68</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Chr, chromosome; OR, odds ratio; 95% CI, 95% confidence interval.

OR is given for a poor histologic response; OR and 95% CI are reported for the minor allele, an OR<1 indicates a protective effect of the minor allele and risk of poor histologic response for the major allele. All reported results are adjusted for neoadjuvant treatment with MTX.

**Association analysis of 5-year PFS**

In the discovery stage, multivariable analysis of genetic variants and 5-year PFS, including metastases at diagnosis as covariable, showed significant associations of 27 SNPs in 21 genes in meta-analysis of the results of the Dutch cohorts. Four of these variants (ABCB11 rs7602171, CYP2D6 rs1058164, CYP2F1 rs7246981, and PPARG rs3856806) were excluded after additional inspection of the genotype clusterplots, leading to 23...
significant variants in 17 different genes (Table 3). In both Dutch cohorts, the same direction of effect was observed for all variants (Supplementary Table S3).

Based on our findings in the discovery phase, 20 variants (excluding three markers in LD) were investigated in the validation cohorts. All variants were in HWE (P>0.05) and had an average call rate of 0.97. After association analysis of the validation cohorts similar to the Dutch cohorts, the \textit{ABCB1} rs17064 variant was independently validated in the Spanish cohort (P=0.014). The \textit{ABCC5} rs562 was significantly associated with 5-year PFS in the English cohort; however, with opposite direction of effect compared to the Dutch cohorts. Although we realized after analysis of the data that the English cohort was not suitable for validation due to its selected population combined with small sample size, we chose to follow the original analysis plan for analysis in the individual cohorts. Furthermore, as the 5-year PFS in the English cohort was substantially different from the Dutch and Spanish cohorts, we carried out a meta-analysis including only the Dutch and Spanish cohorts, which showed inferior 5-year PFS in the Spanish cohort (HR=1.89, P=0.001, after correction for multiple testing \(P_{\text{corrected}}=0.017\)), with a stronger association compared to the meta-analysis of the Dutch cohorts alone. Seven other variants (\textit{ABCC4} rs3742106, \textit{ABCC5} rs562 and rs939336, \textit{ABCG1} rs3788007, \textit{SLC22A5} rs274548 and rs274558, \textit{SLC22A14} rs171248) were also significantly associated with 5-year PFS, with stronger associations than observed in the meta-analysis including the Dutch cohorts only, though these associations did not survive correction for multiple testing. Seven of the eight variants remained significant upon addition of age at diagnosis as a covariate in the analysis, and the other variant, \textit{SLC22A14} rs171248, remained marginally significant (Supplementary Table S3). Addition of the English cohort to the meta-analysis (correcting for metastases only) showed similar results; only two of the eight variants, both in \textit{ABCC5}, were not significantly associated with 5-year PFS, both variants showed large heterogeneity (I²>50) (Supplementary Table S3).

A genetic risk score for PFS was calculated on the basis of the eight variants associated with 5-year PFS in meta-analysis of the Dutch and Spanish cohorts (Figure 1). Patients were grouped by each two consecutive risk scores; patients with ten or more unfavorable alleles were grouped together to ensure sufficient sample size for analysis. Patients with 2-3 risk alleles showed a 5-year PFS of 86.5% compared to a 5-year PFS of 65.8% in patients carrying 4-5 risk alleles, 71.1% with 6-7 alleles, 45.2% with 8-9 alleles, and 34.9% with 10-14 risk alleles (overall P<0.001).

### Table 3. Meta-analysis results of 5-year PFS association analysis in discovery stage.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene</th>
<th>Chr</th>
<th>Minor allele</th>
<th>MAF</th>
<th>HR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs171248</td>
<td>SLC22A14</td>
<td>3</td>
<td>C</td>
<td>0.456</td>
<td>1.45</td>
<td>1.11-1.91</td>
<td>0.007</td>
</tr>
<tr>
<td>rs6441606</td>
<td>CST5</td>
<td>16</td>
<td>C</td>
<td>0.311</td>
<td>0.68</td>
<td>0.50-0.92</td>
<td>0.012</td>
</tr>
<tr>
<td>rs939336</td>
<td>ABCC5</td>
<td>3</td>
<td>T</td>
<td>0.399</td>
<td>1.38</td>
<td>1.07-1.78</td>
<td>0.014</td>
</tr>
<tr>
<td>rs17064</td>
<td>ABCB1</td>
<td>7</td>
<td>T</td>
<td>0.671</td>
<td>1.71</td>
<td>1.11-2.62</td>
<td>0.015</td>
</tr>
<tr>
<td>rs562</td>
<td>ABCC5</td>
<td>3</td>
<td>A</td>
<td>0.472</td>
<td>1.39</td>
<td>1.06-1.84</td>
<td>0.018</td>
</tr>
<tr>
<td>rs149738</td>
<td>SLC22A14</td>
<td>3</td>
<td>C</td>
<td>0.451</td>
<td>1.38</td>
<td>1.05-1.81</td>
<td>0.020</td>
</tr>
<tr>
<td>rs7636919</td>
<td>ABCC5</td>
<td>3</td>
<td>G</td>
<td>0.408</td>
<td>0.73</td>
<td>0.56-0.95</td>
<td>0.021</td>
</tr>
<tr>
<td>rs1695505</td>
<td>ABCG4</td>
<td>7</td>
<td>T</td>
<td>0.221</td>
<td>1.40</td>
<td>1.04-1.87</td>
<td>0.022</td>
</tr>
<tr>
<td>rs1858744</td>
<td>SLC22A14</td>
<td>3</td>
<td>C</td>
<td>0.448</td>
<td>1.36</td>
<td>1.04-1.79</td>
<td>0.027</td>
</tr>
<tr>
<td>rs1364912</td>
<td>CYP3A4</td>
<td>16</td>
<td>T</td>
<td>0.083</td>
<td>0.54</td>
<td>0.31-0.94</td>
<td>0.010</td>
</tr>
<tr>
<td>rs2266762</td>
<td>FMO3</td>
<td>1</td>
<td>A</td>
<td>0.427</td>
<td>1.35</td>
<td>1.03-1.76</td>
<td>0.038</td>
</tr>
<tr>
<td>rs274558</td>
<td>SLC22A5</td>
<td>5</td>
<td>G</td>
<td>0.410</td>
<td>0.75</td>
<td>0.57-0.98</td>
<td>0.036</td>
</tr>
<tr>
<td>rs1822021</td>
<td>SULT1E1</td>
<td>1</td>
<td>A</td>
<td>0.383</td>
<td>1.32</td>
<td>1.02-1.72</td>
<td>0.016</td>
</tr>
<tr>
<td>rs274548</td>
<td>SLC22A5</td>
<td>5</td>
<td>T</td>
<td>0.146</td>
<td>1.40</td>
<td>1.02-1.75</td>
<td>0.017</td>
</tr>
<tr>
<td>rs3742106</td>
<td>ABC4</td>
<td>13</td>
<td>G</td>
<td>0.413</td>
<td>1.34</td>
<td>1.01-1.78</td>
<td>0.041</td>
</tr>
<tr>
<td>rs3788007</td>
<td>ABCG1</td>
<td>21</td>
<td>A</td>
<td>0.196</td>
<td>0.68</td>
<td>0.47-0.99</td>
<td>0.042</td>
</tr>
<tr>
<td>rs1902033</td>
<td>GSTM1</td>
<td>4</td>
<td>G</td>
<td>0.489</td>
<td>0.74</td>
<td>0.55-0.99</td>
<td>0.044</td>
</tr>
<tr>
<td>rs8231712</td>
<td>SULT1E1</td>
<td>4</td>
<td>G</td>
<td>0.141</td>
<td>0.65</td>
<td>0.43-0.99</td>
<td>0.045</td>
</tr>
<tr>
<td>rs7677313</td>
<td>CYP3A4</td>
<td>6</td>
<td>A</td>
<td>0.265</td>
<td>1.33</td>
<td>1.00-1.77</td>
<td>0.047</td>
</tr>
<tr>
<td>rs2780752</td>
<td>CST5</td>
<td>16</td>
<td>T</td>
<td>0.106</td>
<td>0.59</td>
<td>0.35-0.99</td>
<td>0.047</td>
</tr>
<tr>
<td>rs2273657</td>
<td>ABC2</td>
<td>10</td>
<td>A</td>
<td>0.213</td>
<td>1.38</td>
<td>1.00-1.90</td>
<td>0.047</td>
</tr>
<tr>
<td>rs2268877</td>
<td>SLC7A6</td>
<td>14</td>
<td>C</td>
<td>0.214</td>
<td>1.38</td>
<td>1.00-1.90</td>
<td>0.048</td>
</tr>
<tr>
<td>rs9381468</td>
<td>SLC25A27</td>
<td>6</td>
<td>T</td>
<td>0.491</td>
<td>1.32</td>
<td>1.00-1.74</td>
<td>0.059</td>
</tr>
</tbody>
</table>

Chr, chromosome; MAF, minor allele frequency; HR, hazard ratio; 95% CI, 95% confidence interval. HR and 95% CI are reported for the minor allele, a HR <1 indicates a protective effect of the minor allele and risk of poor histologic response for the major allele. All reported results are adjusted for the presence of metastases at diagnosis.

MAF is reported for the total of 283 patients included in the meta-analysis of the two Dutch cohorts. Variants within \textit{CHST5} and \textit{SLC22A14} are in linkage disequilibrium with r²>0.8.
Table 4. Association analysis results of 5-year PFS in validation stage.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene</th>
<th>Chr</th>
<th>Minor allele</th>
<th>Spanish cohort HR</th>
<th>95% CI</th>
<th>P</th>
<th>English cohort HR</th>
<th>95% CI</th>
<th>P</th>
<th>Meta-analysis Dutch and Spanish cohorts HR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs17064</td>
<td>ABCB1</td>
<td>7</td>
<td>T</td>
<td>2.59</td>
<td>1.21-5.54</td>
<td>0.014</td>
<td>0.55</td>
<td>0.07-4.43</td>
<td>0.577</td>
<td>1.89</td>
<td>1.30-2.74</td>
<td>0.001</td>
</tr>
<tr>
<td>rs562</td>
<td>ABCC5</td>
<td>3</td>
<td>A</td>
<td>1.59</td>
<td>0.96-2.64</td>
<td>0.070</td>
<td>0.38</td>
<td>0.17-0.89</td>
<td>0.026</td>
<td>1.44</td>
<td>1.13-1.83</td>
<td>0.003</td>
</tr>
<tr>
<td>rs174206</td>
<td>ABCC4</td>
<td>13</td>
<td>G</td>
<td>1.56</td>
<td>0.38-2.47</td>
<td>0.058</td>
<td>0.80</td>
<td>0.36-1.78</td>
<td>0.589</td>
<td>1.40</td>
<td>1.10-1.78</td>
<td>0.006</td>
</tr>
<tr>
<td>rs274558</td>
<td>SLC22A5</td>
<td>5</td>
<td>G</td>
<td>0.60</td>
<td>0.34-1.06</td>
<td>0.079</td>
<td>0.86</td>
<td>0.42-1.76</td>
<td>0.679</td>
<td>0.72</td>
<td>0.56-0.92</td>
<td>0.008</td>
</tr>
<tr>
<td>rs933364</td>
<td>ABCC5</td>
<td>3</td>
<td>T</td>
<td>1.27</td>
<td>0.77-2.10</td>
<td>0.354</td>
<td>0.44</td>
<td>0.18-1.06</td>
<td>0.068</td>
<td>1.36</td>
<td>1.08-1.70</td>
<td>0.009</td>
</tr>
<tr>
<td>rs3888037</td>
<td>ARFGT</td>
<td>21</td>
<td>A</td>
<td>0.50</td>
<td>0.23-1.11</td>
<td>0.067</td>
<td>1.54</td>
<td>0.61-3.89</td>
<td>0.357</td>
<td>0.64</td>
<td>0.46-0.90</td>
<td>0.010</td>
</tr>
<tr>
<td>rs1277388</td>
<td>SLCO4A5</td>
<td>5</td>
<td>T</td>
<td>1.54</td>
<td>0.72-2.38</td>
<td>0.268</td>
<td>1.60</td>
<td>0.77-3.33</td>
<td>0.210</td>
<td>1.49</td>
<td>1.07-2.08</td>
<td>0.018</td>
</tr>
<tr>
<td>rs173124</td>
<td>SLC22A4</td>
<td>3</td>
<td>C</td>
<td>0.90</td>
<td>0.56-1.45</td>
<td>0.467</td>
<td>1.00</td>
<td>0.41-2.45</td>
<td>0.991</td>
<td>1.28</td>
<td>1.02-1.63</td>
<td>0.035</td>
</tr>
<tr>
<td>rs3822173</td>
<td>SULT1E1</td>
<td>21</td>
<td>G</td>
<td>0.88</td>
<td>0.42-1.84</td>
<td>0.736</td>
<td>1.76</td>
<td>0.52-5.27</td>
<td>0.388</td>
<td>0.70</td>
<td>0.49-1.01</td>
<td>0.056</td>
</tr>
<tr>
<td>rs1095055</td>
<td>ABCB4</td>
<td>7</td>
<td>T</td>
<td>0.88</td>
<td>0.46-1.68</td>
<td>0.700</td>
<td>0.74</td>
<td>0.24-2.25</td>
<td>0.597</td>
<td>1.29</td>
<td>0.99-1.68</td>
<td>0.060</td>
</tr>
<tr>
<td>rs184202</td>
<td>CDA</td>
<td>10</td>
<td>A</td>
<td>1.00</td>
<td>0.64-1.57</td>
<td>0.995</td>
<td>1.51</td>
<td>0.72-3.18</td>
<td>0.280</td>
<td>1.23</td>
<td>0.89-1.55</td>
<td>0.071</td>
</tr>
<tr>
<td>rs2276497</td>
<td>ARRC5</td>
<td>3</td>
<td>A</td>
<td>1.03</td>
<td>0.57-1.85</td>
<td>0.915</td>
<td>0.91</td>
<td>0.27-3.05</td>
<td>0.876</td>
<td>1.29</td>
<td>0.89-1.71</td>
<td>0.074</td>
</tr>
<tr>
<td>rs7664910</td>
<td>SLC7A4</td>
<td>14</td>
<td>C</td>
<td>1.18</td>
<td>0.70-2.00</td>
<td>0.532</td>
<td>0.75</td>
<td>0.27-2.08</td>
<td>0.582</td>
<td>0.81</td>
<td>0.64-1.02</td>
<td>0.077</td>
</tr>
<tr>
<td>rs2288072</td>
<td>SLC7A4</td>
<td>14</td>
<td>C</td>
<td>0.77</td>
<td>0.41-1.45</td>
<td>0.413</td>
<td>1.36</td>
<td>0.43-3.91</td>
<td>0.644</td>
<td>1.23</td>
<td>0.92-1.63</td>
<td>0.163</td>
</tr>
<tr>
<td>rs2266752</td>
<td>HMOX1</td>
<td>1</td>
<td>A</td>
<td>0.84</td>
<td>0.50-1.43</td>
<td>0.523</td>
<td>0.61</td>
<td>0.23-1.58</td>
<td>0.406</td>
<td>1.20*</td>
<td>0.81-1.78</td>
<td>0.35</td>
</tr>
<tr>
<td>rs5984186</td>
<td>SLC25A27</td>
<td>6</td>
<td>G</td>
<td>0.79</td>
<td>0.47-1.32</td>
<td>0.370</td>
<td>1.48</td>
<td>0.73-2.99</td>
<td>0.278</td>
<td>1.17*</td>
<td>0.68-1.87</td>
<td>0.45</td>
</tr>
<tr>
<td>rs1902023</td>
<td>UGT2B15</td>
<td>4</td>
<td>G</td>
<td>1.10</td>
<td>0.80-1.33</td>
<td>0.267</td>
<td>0.62</td>
<td>0.28-1.38</td>
<td>0.242</td>
<td>0.89*</td>
<td>0.61-1.28</td>
<td>0.52</td>
</tr>
<tr>
<td>rs7367131</td>
<td>OPN3M1</td>
<td>6</td>
<td>A</td>
<td>0.72</td>
<td>0.42-1.23</td>
<td>0.224</td>
<td>1.06</td>
<td>0.47-2.37</td>
<td>0.897</td>
<td>1.13*</td>
<td>0.75-1.71</td>
<td>0.56</td>
</tr>
<tr>
<td>rs1364182</td>
<td>GPAP</td>
<td>16</td>
<td>T</td>
<td>1.60</td>
<td>0.80-3.19</td>
<td>0.184</td>
<td>0.51</td>
<td>0.14-1.89</td>
<td>0.312</td>
<td>0.82*</td>
<td>0.38-1.76</td>
<td>0.61</td>
</tr>
<tr>
<td>rs6414086</td>
<td>CNST5</td>
<td>16</td>
<td>C</td>
<td>1.64</td>
<td>1.02-2.64</td>
<td>0.042</td>
<td>0.48</td>
<td>0.15-1.51</td>
<td>0.208</td>
<td>0.90*</td>
<td>0.52-1.58</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Chr, chromosome; HR, hazard ratio; 95% CI, 95% confidence interval. HR and 95% CI are reported for the minor allele, a HR <1 indicates a protective effect of the minor allele and risk of poor histologic response for the major allele. All reported results are adjusted for the presence of metastases at diagnosis.

I*2>50.
aResults for the English and Spanish cohort are given for rs939338: the genotyped variant in these cohorts.

Figure 1. Five-year PFS based on GRS in the Dutch and Spanish cohorts combined (N=364). The GRS is based on eight markers significantly associated with 5-year PFS in meta-analysis of Dutch and Spanish cohorts. Patients with 2-3 risk alleles showed a 5-year PFS of 86.5% compared to 65.8% with 4-5 risk alleles, 71.1% with 6-7 alleles, 45.2% with 8-9 alleles, and 34.9% with 10-14 risk alleles.

Discussion

In recent years, pharmaco genetic studies have explored the opportunity to distinguish osteosarcoma patients prone to have a poor response to chemotherapy from those who will benefit from the treatment, by studying candidate genes in the metabolic pathways of the chemotherapeutics used in antitumor treatment. In the present study, we have taken the next step by interrogating a comprehensive set of drug metabolism and transporter genes to detect associations with response to chemotherapy. We have identified eight genetic variants associated with 5-year PFS, of which an ABCB1 variant was independently validated in one of the validation cohorts. Seven other variants were significantly associated with 5-year PFS in a combined meta-analysis of discovery and validation cohorts.

In addition to 5-year PFS, the histologic response was included as endpoint in our study, as in clinical practice this is used as a marker of the response to neoadjuvant chemotherapy, and as the basis for stratification in past trials, based on its presumed predictive capacity of outcome. Remarkably, in only one of the four cohorts (Dutch cohort II) included in the present study, the histologic response was associated with 5-year PFS. Moreover, despite significant differences in poor responders between the two Dutch cohorts, this did not translate into a difference in PFS. This is in line with trials of intensified treatment.
leading to better histologic response but not to significantly improved outcome [2,3]. In the genetic association analysis, none of the markers most significantly associated in the Dutch cohorts could be validated. Conflicting results were obtained on the T allele of the ABCC2 rs171620 variant, which had been previously shown to be associated with poor histologic response in the English cohort [4] and confirmed in our analysis of the same cohort, but which showed a protective effect in the Dutch cohorts. As analysis of the Spanish cohort did not provide significant results, the influence of this variant on the histologic response remains doubtful. A possible explanation for the discrepancy between the results of the Dutch and English cohorts is the large difference in the proportion of poor histologic responders. The extremely small proportion of poor responders in the English cohort might be attributed to the absence of patients who had died before the time of retrospective patient inclusion, of whom a substantial proportion is expected to have had a poor histologic response, even though the link between the histologic response and survival may not be absolute as indicated above. The 100% neoadjuvant MTX treatment in the English cohort, compared to lower percentages in the other cohorts, may also have influenced the histologic response, as neoadjuvant MTX treatment was related to the histologic response in Dutch cohort 1.

In the English cohort, not only a poor histologic response was less frequently observed than expected, also the 5-year PFS (and OS) was substantially higher than in the other cohorts, which we consider to have the same cause. Unfortunately, this may indicate that a different sample of patients was selected, and together with the relatively small sample size, the results of the English cohort were considered inconclusive for this study. For this reason, the GRS analysis was only carried out with the combined Dutch and Spanish cohorts. Aside from the expected overrepresentation of survivors, the English cohort also differs from the other cohorts with respect to ethnic background, MTX treatment, and time frame of inclusion. However, the percentage of patients with overall MTX treatment in the English cohort, compared to the Spanish cohort. Furthermore, (overall) survival differences between ethnicities, if any, are directed towards inferior survival in black patients [12], and improvement in survival due to for example better supportive care in the more recent time frame is often relatively subtle. Hence, these factors are not expected to have substantial impact on the 5-year PFS such as observed in the English cohort.

The GRS analysis included eight genetic variants in six different genes associated with 5-year PFS in meta-analysis of the Dutch and Spanish cohorts, to evaluate the combined effect on PFS. Hence, of two genes (ABCC5 and SLC22A5), two variants per gene were included in the analysis. As the LD between the variants within these genes was only partial (r2<0.8), inclusion of both variants per gene was not expected to have led to overestimation, as each allele may still have an additional effect.

The rs17064 variant independently validated in relation to PFS is located in the ABCB1 (ATP-binding cassette, sub-family B member 1) gene. This gene encodes ABCB1 (P-glycoprotein), an efflux transporter involved in elimination through bile and the kidneys; and MTX, doxorubicin and vincristine are among its known substrates [13]. At the tumor level, this transporter has received attention as previous studies have indicated that ABCB1 overexpression in osteosarcoma is related to doxorubicin resistance and treatment response [14-16], and an ABCB1/ABCC1 inhibitor has been shown to revert ABCB1/ABCC1-mediated doxorubicin resistance of osteosarcoma in vitro [17]. Germline pharmacogenetic studies in osteosarcoma have reported conflicting results regarding the influence of ABCB1 variants on treatment response: other variants of ABCB1, including rs1128503 and rs1045642, have been reported in association with histologic response and/or survival [5,18], but have shown opposite results [19,20] or lack of association [4] in other studies. These variants are both included on the DMET array, but were also not significantly associated with either histologic response or 5-year PFS in the present study. The variant identified in the present study has not been studied before in osteosarcoma, but in an ovarian cancer pharmacogenetic study in which other ABCB1 variants were associated with outcome after platinum-based treatment, though not rs17064 [21]. The functional consequence of the variant is not known, but as the variant is located in the 3′UTR of the gene, it might have an effect on expression, or alternatively it may not be the causal variant itself but be linked to a functional variant. Another transporter gene that was significant in the PFS analysis is ABCC5 (ATP-binding cassette sub-family C member 5), which was also identified in our pathway-based study [6]. Of both the rs562 variant in the 3′UTR and the synonymous rs939336 variant, an eQTL study has linked the risk alleles to increased ABCC5 expression [22], which on its turn has been implicated in cisplatin and doxorubicin resistance [23,24]. As the synonymous rs939336 variant covered by the DMET array is almost completely linked to the rs939338 variant detected in our previous study, this finding may in fact be considered as an internal validation with another genotyping platform. We also identified an intronic variant in ABCG1 (ATP-binding cassette sub-family G member 1), which encodes an efflux transporter of cholesterol in various cell types [25]. Interestingly, ABCG1 variants have been linked to overall survival after platinum-based chemotherapy in non-small cell lung cancer [26]. Another association found was of an 3′UTR variant in ABCG4 (ATP-binding cassette sub-family G member 4). This gene encodes an efflux transporter known to function in brain, liver and kidneys. ABCG4 has been implicated in cisplatin-resistance in gastric cancer [27]. Moreover, overexpression of ABCG4 has been correlated with multidrug resistant osteosarcoma [28]. Furthermore, two other genes, SLC22A5 (Solute carrier family 22 member 5) and SLC22A14 (Solute carrier family 22 member 14) were found in relation to PFS in the present study that encode transporters located in the kidneys. SLC22A5 is involved in the reabsorption of carnitine and has been implicated in cisplatin-induced nephrotoxicity [29,30]. SLC22A14 (or OCTL2) is an organic cation transporter which has not been previously linked to transport of anticancer drugs [31].
All genes identified in the current study are coding for drug transporters; although a large number of phase I and phase II metabolic enzymes were covered by the DMET array, none of these were significant in our study. Although previous candidate gene or pathway-based pharmacogenetic studies in osteosarcoma have indicated a role of variation in various glutathione-S-transferase genes in the treatment response, with mostly consistent findings for the GSTP1 rs1695 variant which is also included on the DMET array, we could not confirm these findings [4,18,19,32-35].

To our knowledge, this is the largest pharmacogenetic study in osteosarcoma to date with regard to the number of included patients, inclusion of validation cohorts, and the number of investigated genes and markers. However, we acknowledge that the sample size is still relatively low for detection of genetic associations, reflecting the low incidence of osteosarcoma. Although significant results were obtained, we may have missed other associations, as differences between good and poor responders may have been limited by the low number of patients. In the context of genetic association analysis with relatively small cohorts, we have applied less stringent procedures for selection of variants to take forward to validation, i.e. we did not correct for multiple testing, and we only investigated the potential influence of age as additional covariable in the PFS analysis after validation of markers. In addition, the Dutch cohorts were heterogeneous regarding treatment protocols. Part of the patients received only two drugs (cisplatin and doxorubicin), whereas other patients also received additional drugs (mostly only MTX), which might give a more favorable outcome. However, as there was no effect of the presence of MTX in the treatment regimen on PFS in our cohorts, the influence of the differences in treatment regimens on the results is probably limited. Furthermore, both the genetic background of the tumor which might also influence the chemotherapy response, and tumor-specific mutations in drug transporter genes have not been taken into account in the study. The latter may in general not be of major impact on transporter genes, as DNA derived from tumor and blood or saliva has been shown to provide highly concordant genotypes of variants involved in drug metabolism and transport [36]. Nonetheless, it would be very interesting to compare that at the tumor level in osteosarcoma, as has been done ever since the early nineties and has been published as predictive factor already at that time [14].

For patients carrying risk variants, the question arises as to how this could be applied for treatment optimization in the future. In general, for variants in genes related to pharmacokinetics, increasing the dose might give a more favorable outcome. However, this is complicated by the risk of toxicities of chemotherapeutic treatment for osteosarcoma. Moreover, as drug transporters such as ABCB1 are not only present in typical drug processing organs but also in the tumor itself, prediction of the effect of a genetic variant may also require characterization at the tumor level. Alternatively, a poor predictive profile may facilitate identification of patients for whom alternative treatment, once it is available, is preferred.

In conclusion, we have contributed to the evidence for a role of ABCB1 variants in the response to cisplatin and doxorubicin-based chemotherapy, and have revealed novel associations of other transporters not known to be directly related to the chemotherapeutics used in the treatment of osteosarcoma. Together with the genetic variants previously identified in our pathway-based study, the associations of the present study warrant further confirmation in homogeneous large cohorts, and functional validation to provide additional evidence and increase our understanding of the treatment response in osteosarcoma. In the future, a set of validated markers might provide a useful tool for upfront risk stratification of patients with osteosarcoma to eventually optimize current therapy.
References


SNPs in ADME genes and osteosarcoma treatment response


Supplementary data

Table S1. Demographic data of eligible patients for PFS analysis in the discovery and validation cohorts.

<table>
<thead>
<tr>
<th></th>
<th>Dutch cohort I (N=118)</th>
<th>Dutch cohort II (N=165)</th>
<th>Spanish cohort (N=95)</th>
<th>English cohort (N=51)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis, median (range)</td>
<td>15.6 (5.6-41.1)</td>
<td>17.1 (3.4-45.8)</td>
<td>14.0 (1.9-34.6)</td>
<td>17.0 (10.0-39.0)</td>
</tr>
<tr>
<td>Male gender, n (%)</td>
<td>64 (54.2%)</td>
<td>93 (56.4%)</td>
<td>53 (55.8%)</td>
<td>29 (56.9%)</td>
</tr>
<tr>
<td>Axial tumor, n (%)</td>
<td>5 (4.2%)</td>
<td>7 (4.2%)</td>
<td>2 (2.1%)</td>
<td>3 (5.9%)</td>
</tr>
<tr>
<td>Primary metastases, n (%)</td>
<td>22 (18.6%)</td>
<td>24 (14.5%)</td>
<td>17 (17.9%)</td>
<td>6 (11.8%)</td>
</tr>
<tr>
<td>MTX total, n (%)</td>
<td>77 (65.3%)</td>
<td>76 (46.1%)</td>
<td>84 (95.5%)</td>
<td>51 (100%)</td>
</tr>
<tr>
<td>MTX neoadjuvant, n (%)</td>
<td>65 (55.1%)</td>
<td>70 (42.4%)</td>
<td>30 (35.3%)</td>
<td>50 (100%)</td>
</tr>
<tr>
<td>Poor histologic response, n (%)</td>
<td>80 (66.6%)</td>
<td>75 (48.7%)</td>
<td>49 (55.1%)</td>
<td>11 (22.0%)</td>
</tr>
<tr>
<td>5-year PFS</td>
<td>60.5%</td>
<td>60.8%</td>
<td>58.2%</td>
<td>74.5%</td>
</tr>
<tr>
<td>5-year OS</td>
<td>73.6%</td>
<td>75.4%</td>
<td>83.6%</td>
<td>88.2%</td>
</tr>
</tbody>
</table>

MTX, methotrexate; PFS, progression-free survival; OS, overall survival.

a Dutch cohorts I and II significantly differed regarding age at diagnosis (P=0.049), MTX treatment total (P=0.001) and neoadjuvant (P=0.036), and poor histologic response (P=0.001).

b Ethnicities English cohort: Caucasian N=36 (70.6%), Afro-Caribbean N=8 (15.7%), Indian/Asian N=7 (13.3%); Spanish cohort: Caucasian N=90 (94.7%), Latin American N=4 (4.2%), Romani N=1 (1.1%).
Table S2. Complete association analysis results of histological response in the Dutch cohorts.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene</th>
<th>Chr</th>
<th>Minor allele</th>
<th>Dutch cohort I OR 95% CI</th>
<th>P</th>
<th>Dutch cohort II OR 95% CI</th>
<th>P</th>
<th>Meta-analysis Dutch cohorts OR 95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs717620</td>
<td>ABC2</td>
<td>10</td>
<td>T</td>
<td>0.66 0.33-1.33</td>
<td>0.243</td>
<td>0.36 0.19-0.68</td>
<td>0.002</td>
<td>0.67 0.39-0.76</td>
<td>0.002</td>
</tr>
<tr>
<td>rs9381468</td>
<td>SLC25A27</td>
<td>6</td>
<td>T</td>
<td>2.13 1.17-3.87</td>
<td>0.013</td>
<td>1.55 0.98-2.46</td>
<td>0.064</td>
<td>1.75 1.21-2.52</td>
<td>0.003</td>
</tr>
<tr>
<td>rs284737</td>
<td>VKORC1</td>
<td>16</td>
<td>G</td>
<td>0.53 0.27-1.04</td>
<td>0.064</td>
<td>0.60 0.36-1.00</td>
<td>0.069</td>
<td>0.57 0.38-0.86</td>
<td>0.007</td>
</tr>
<tr>
<td>rs8038094</td>
<td>VKORC1</td>
<td>16</td>
<td>C</td>
<td>0.60 0.33-1.09</td>
<td>0.093</td>
<td>0.60 0.37-0.98</td>
<td>0.041</td>
<td>0.60 0.41-0.88</td>
<td>0.008</td>
</tr>
<tr>
<td>rs249221</td>
<td>ABC2</td>
<td>16</td>
<td>C</td>
<td>2.18 1.05-4.53</td>
<td>0.038</td>
<td>1.46 0.88-2.43</td>
<td>0.147</td>
<td>1.66 1.09-2.52</td>
<td>0.017</td>
</tr>
<tr>
<td>rs2011537</td>
<td>SLC25A27</td>
<td>13</td>
<td>T</td>
<td>0.70 0.38-1.31</td>
<td>0.264</td>
<td>0.61 0.38-0.97</td>
<td>0.037</td>
<td>0.64 0.44-0.93</td>
<td>0.019</td>
</tr>
<tr>
<td>rs9934038</td>
<td>VKORC1</td>
<td>16</td>
<td>T</td>
<td>0.64 0.35-1.15</td>
<td>0.135</td>
<td>0.64 0.39-1.04</td>
<td>0.073</td>
<td>0.64 0.44-0.93</td>
<td>0.020</td>
</tr>
<tr>
<td>rs11572080</td>
<td>CPY2C8</td>
<td>10</td>
<td>A</td>
<td>0.43 0.17-1.06</td>
<td>0.066</td>
<td>0.53 0.19-1.46</td>
<td>0.219</td>
<td>0.47 0.24-0.92</td>
<td>0.028</td>
</tr>
<tr>
<td>rs2295475</td>
<td>ADM</td>
<td>2</td>
<td>T</td>
<td>0.72 0.39-1.32</td>
<td>0.292</td>
<td>0.62 0.38-1.00</td>
<td>0.050</td>
<td>0.66 0.45-0.96</td>
<td>0.029</td>
</tr>
<tr>
<td>rs7708472</td>
<td>VKORC1</td>
<td>16</td>
<td>T</td>
<td>1.68 0.79-3.59</td>
<td>0.179</td>
<td>1.69 0.92-3.11</td>
<td>0.090</td>
<td>1.69 1.05-2.72</td>
<td>0.030</td>
</tr>
<tr>
<td>rs1801265</td>
<td>DPD</td>
<td>1</td>
<td>C</td>
<td>0.78 0.41-1.49</td>
<td>0.454</td>
<td>0.57 0.34-0.96</td>
<td>0.034</td>
<td>0.65 0.43-0.97</td>
<td>0.034</td>
</tr>
<tr>
<td>rs2808840</td>
<td>CY2C18</td>
<td>10</td>
<td>T</td>
<td>0.51 0.27-0.97</td>
<td>0.042</td>
<td>0.78 0.48-1.27</td>
<td>0.315</td>
<td>0.67 0.45-0.99</td>
<td>0.042</td>
</tr>
<tr>
<td>rs202861</td>
<td>FMO2</td>
<td>1</td>
<td>G</td>
<td>1.93 1.03-3.60</td>
<td>0.041</td>
<td>1.27 0.81-2.00</td>
<td>0.299</td>
<td>1.46 1.01-2.09</td>
<td>0.042</td>
</tr>
<tr>
<td>rs740666</td>
<td>ABC2</td>
<td>10</td>
<td>T</td>
<td>0.83 0.44-1.54</td>
<td>0.450</td>
<td>0.61 0.38-0.98</td>
<td>0.040</td>
<td>0.68 0.47-0.99</td>
<td>0.046</td>
</tr>
<tr>
<td>rs1442477</td>
<td>ADH1</td>
<td>4</td>
<td>A</td>
<td>0.38 0.09-1.49</td>
<td>0.053</td>
<td>0.24 0.05-1.16</td>
<td>0.075</td>
<td>0.28 0.08-0.99</td>
<td>0.048</td>
</tr>
</tbody>
</table>

Chr, chromosome; OR, odds ratio; 95% CI, 95% confidence interval.
OR is given for a poor histologic response; OR and 95% CI are reported for the minor allele, an OR<1 indicates a protective effect of the minor allele and risk of poor histologic response for the major allele.
All reported results are adjusted for neoadjuvant treatment with MTX.

Figure S1. Five-year PFS of all patients included in the study for discovery (Dutch cohorts) and validation (English and Spanish cohorts).
<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene</th>
<th>Chr</th>
<th>Minor allele</th>
<th>HR</th>
<th>95% CI</th>
<th>P</th>
<th>HR</th>
<th>95% CI</th>
<th>P</th>
<th>HR</th>
<th>95% CI</th>
<th>P</th>
<th>HR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs17064</td>
<td>ABCB1</td>
<td>7</td>
<td>T</td>
<td>1.69</td>
<td>0.78-3.68</td>
<td>0.184</td>
<td>1.71</td>
<td>1.02-2.87</td>
<td>0.041</td>
<td>1.82</td>
<td>1.26-2.63</td>
<td>0.001</td>
<td>1.91</td>
<td>1.31-2.80</td>
</tr>
<tr>
<td>rs562</td>
<td>ABCC5</td>
<td>3</td>
<td>A</td>
<td>1.79</td>
<td>1.14-2.81</td>
<td>0.012</td>
<td>1.20</td>
<td>0.85-1.71</td>
<td>0.296</td>
<td>1.19*</td>
<td>0.74-1.91</td>
<td>0.48</td>
<td>1.42</td>
<td>1.11-1.82</td>
</tr>
<tr>
<td>rs3747086</td>
<td>ABC4</td>
<td>13</td>
<td>G</td>
<td>1.21</td>
<td>0.76-1.92</td>
<td>0.407</td>
<td>1.43</td>
<td>1.00-2.03</td>
<td>0.049</td>
<td>1.33</td>
<td>1.06-1.68</td>
<td>0.014</td>
<td>1.39</td>
<td>1.09-1.77</td>
</tr>
<tr>
<td>rs7244558</td>
<td>SC22A5</td>
<td>5</td>
<td>G</td>
<td>0.85</td>
<td>0.56-1.30</td>
<td>0.456</td>
<td>0.68</td>
<td>0.46-0.97</td>
<td>0.014</td>
<td>0.73</td>
<td>0.58-0.92</td>
<td>0.008</td>
<td>0.74</td>
<td>0.58-0.95</td>
</tr>
<tr>
<td>rs93936</td>
<td>ABC5</td>
<td>3</td>
<td>T</td>
<td>1.68</td>
<td>1.10-2.66</td>
<td>0.016</td>
<td>1.23</td>
<td>0.89-1.70</td>
<td>0.215</td>
<td>1.19*</td>
<td>0.81-1.73</td>
<td>0.37</td>
<td>1.37</td>
<td>1.09-1.73</td>
</tr>
<tr>
<td>rs5960007</td>
<td>ABCG2</td>
<td>21</td>
<td>G</td>
<td>0.85</td>
<td>0.56-1.30</td>
<td>0.456</td>
<td>0.68</td>
<td>0.46-0.97</td>
<td>0.014</td>
<td>0.73</td>
<td>0.58-0.92</td>
<td>0.008</td>
<td>0.74</td>
<td>0.58-0.95</td>
</tr>
<tr>
<td>rs724548</td>
<td>SC22A5</td>
<td>5</td>
<td>T</td>
<td>1.37</td>
<td>0.77-2.42</td>
<td>0.284</td>
<td>1.57</td>
<td>0.97-2.55</td>
<td>0.067</td>
<td>1.51</td>
<td>1.12-2.04</td>
<td>0.008</td>
<td>1.41</td>
<td>1.01-1.98</td>
</tr>
</tbody>
</table>

Chr, chromosome; HR, hazard ratio; 95% CI, 95% confidence interval, NA, not analyzed.

HR and 95% CI are reported for the minor allele, a HR <1 indicates a protective effect of the minor allele and risk of poor histologic response for the major allele. All reported results are adjusted for the presence of metastases at diagnosis, except for meta-analysis results of the Dutch and Spanish cohorts which included correction for both metastases and age at diagnosis, analyzed for the eight variant significantly associated with 5-year PFS with correction for only metastases. Results are sorted as in Table 4 for comparison.

*P<0.05.
Chapter 7

Pharmacogenetics of chemotherapy response in osteosarcoma: a genetic variant in SLC7A8 is associated with progressive disease


* These authors contributed equally

Submitted
Abstract

Despite (neo)adjuvant chemotherapy in primary osteosarcoma or first-line chemotherapy in patients presenting with metastatic disease, some patients progress already during first-line systemic treatment and thus have a very poor prognosis. Here we investigated whether patients with an inadequate response to treatment have a distinctive pharmacogenetic profile, by applying a large scale screening including 1,936 genetic markers in 231 drug metabolism and transporter genes.

Germline DNA from 287 Dutch high-grade osteosarcoma patients treated with cisplatin and doxorubicin-based chemotherapy was genotyped using the DMET Plus array. Associations between genetic variants and progressive disease (primary tumor and/or metastasis growth or formation up to 3 months after end of adjuvant chemotherapy or first-line treatment in case of primary metastatic disease; and/or inadequacy to reach complete remission at the end of therapy for primary localized or primary metastatic osteosarcoma) were assessed using logistic regression models. Genetic variants significantly associated with progressive disease were validated in an independent cohort of 146 patients from Spain and England, followed by a second validation using a cohort of 28 patients from Australia.

In multivariable association analyses of genetic variants and progressive disease, adjusted for the presence of primary metastases, sex and age at diagnosis, 10 genetic variants in 6 genes were uniquely associated (P<0.05) with progressive disease. Of these variants, SLC7A8 rs1884545 was independently validated, and showed increased significance in meta-analysis of all cohorts combined (OR 0.22 [0.07-0.63], P=0.005).

Three other variants, CYP8B1 rs6771233, SLC22A2 rs316003, and SLC22A5 rs274548, were also significantly associated in meta-analysis of all cohorts.

We have identified genetic variants specifically associated with an inadequate drug response, defined as progressive disease, in patients with osteosarcoma. This analysis represents the first step towards identifying patients for whom chemotherapeutic treatment is ineffective and should be further explored in additional patient cohorts.

Introduction

Osteosarcoma is a malignant bone tumor that mainly affects children and adolescents. Although the number of people affected by osteosarcoma is low (worldwide 3-4 patients per million), the disease is ranked as one of the most frequent causes of cancer related death in young patients [1]. The disease has a great impact on the patient’s life, as treatment requires an intensive combination of chemotherapy, often disabling surgery, and prolonged periods of rehabilitation. Despite this harsh treatment regimen, some patients fail to respond to the treatment showing no response or even tumor growth during primary treatment. Already for decades the backbone of chemotherapy in the first-line treatment of osteosarcoma consist of a combination containing at least cisplatin and doxorubicin. However, these inadequately responding patients underscore the need for alternative medical treatment options at presentation. Several trials, mainly focusing on osteosarcoma patients with recurrent disease, have indicated that a number of patients (10-45%) show an objective response to second line treatment [2-7]. In addition, although few, new drugs are in the pipeline with promising results for clinical implementation for patients with progressive or recurrent disease. Recognition of patients that do not benefit from current chemotherapy schedules at an early phase in treatment is therefore important.

Although the genetics and biology of the tumor are likely to contribute to the heterogeneous response to treatment, we postulate that germline variants in drug metabolizing enzymes or transporters may also contribute to this observed heterogeneity. Pharmacogenetics holds the promise to discover germline genetic variants predictive of the drug response in individual patients. Most studies aimed at identifying germline genetic variants predictive of treatment outcome in osteosarcoma have used survival as the main clinical endpoint. However, patients with progressive disease, who may have a distinctive pharmacogenetic profile, have not been widely considered as a subgroup. The few studies comparing patients showing a good clinical response (complete and partial responders) to patients with a poor response (stable or progressive disease) investigated only a few candidate genes involved in DNA repair (CCNH, ERCC1/2/5/6, MMS19L and XPC) and GSTP1, a gene involved in detoxification of exogenous and endogenous compounds [8-12]. Significant associations of genetic variants in ERCC2/5, MMS19L, XPC and GSTP1 with clinical response have been identified. However, validation in additional samples is necessary to confirm these associations, because the results presented are often conflicting [10-12], and a more comprehensive screening of genetic variants involved in drug metabolism and transport is warranted.
Therefore, we have performed a large scale screening of 1,936 genetic variants in 231 drug metabolism and transporter genes in a large group of osteosarcoma patients aimed at discovering a relation to an inadequate drug response during first treatment, followed by validation in independent cohorts of patients.

Materials and methods

Patient cohorts

A discovery cohort of 287 osteosarcoma patients was retrospectively collected at the Radboud university medical center (Nijmegen), the University Medical Center of Groningen (Groningen), Leiden University Medical Center (Leiden) and the Emma Children’s Hospital/Academic Medical Center (Amsterdam), The Netherlands. All patients were treated between 1978 and 2013 and the clinical data was retrospectively collected from medical records. Eligibility criteria were: histological diagnosis of primary high-grade osteosarcoma with or without metastatic disease, age ≤45 years, treatment with cisplatin and doxorubicin-based chemotherapy (also neoadjuvant), and self-reported Caucasian ethnicity. Patients were treated either according to institutional standard therapy consisting of cisplatin (maximum cumulative dose 600 mg/m²) and doxorubicin (maximum cumulative dose 450 mg/m²), or according to the standard schedule as given in the EURAMOS-1 trial, which consisted of cisplatin (480 mg/m²) / doxorubicin (450 mg/m²) and additionally high-dose methotrexate (MTX; 144 g/m²), with or without additional ifosfamide/etoposide or interferon-α [13]. The study was approved by the Institutional review board of the Radboud university medical center, and approval for inclusion of patients in other institutions was obtained from institutional ethics committees. All patients and/or parents provided written informed consent.

A combined cohort of 146 high-grade osteosarcoma patients treated with cisplatin and doxorubicin-based chemotherapy from Spain (N=95) and England (N=51) was used for independent validation of the positive findings in the discovery cohort [14,15]. Information on treatment has been reported previously [14,15]. In a second validation phase, a cohort of 28 high-grade osteosarcoma patients, treated with cisplatin and doxorubicin-based chemotherapy from Australia (Sydney Children’s Tumour Bank Network), was included. Patients were treated with a cisplatin cumulative dose of 480 mg/m² or 600 mg/m² and doxorubicin cumulative dose of 450 mg/m². The same inclusion criteria as used in the discovery cohort were applied, with the exception of ethnicity. The study was approved by the local ethics committees of these validation cohorts and written informed consent was obtained from parents and/or patients.

Response definition

The clinical (radiological) response to the treatment was based on imaging results (CT/ MRI/X-ray) reviewed by local expert radiologists. An initial inadequate drug response was operationalized as progressive disease. This was defined as: (1) growth of the primary tumor (>20%) and/or metastases (>20%), or development of new lesions, in the time from start of primary treatment until 3 months after end of adjuvant chemotherapy or end of first-line treatment in case of primary metastatic disease, and/or (2) inadequacy to reach complete remission at the end of (surgical and chemotherapeutic) therapy for primary localized or primary metastatic osteosarcoma. The opposite extremes, patients showing an adequate drug response with no signs of relapse were considered controls. Thus patients with recurrent disease, defined as local or distant relapse from 3 months after end of primary treatment to end of follow-up, were excluded from the analysis.

Genotyping methods

For the discovery cohort, germline DNA was isolated from blood using the QIaamp DNA Blood Midi kit (Qiagen, Venlo, The Netherlands), or from saliva using the Oragene saliva collection kit (DNA Genotek, Kanata, Ontario, Canada) according to the manufacturer’s protocols. From patients who had died before inclusion in this pharmacogenetic study, DNA was isolated from normal formalin-fixed, paraffin-embedded (FFPE) tissue as described previously [16,17]. The DNA samples were genotyped for 1,936 genetic variants using the Affymetrix DMET Plus array according to the manufacturer’s instructions (Affymetrix UK Ltd, High Wycombe, UK). All samples were normalized to 60 ng/μL based on NanoDrop measurements (NanoDrop 2000; Thermo Scientific, Breda, The Netherlands). Genotypes were calculated with DMET console software 1.3 using the Dynamic Genotype Boundaries version 2 algorithm. Variants were excluded from analysis if the genotype cluster plots were considered unreliable, being plots with genotype calling showing merged clusters without distinct cluster boundaries. Additional stringent evaluation of the genotype clustering in combination with expected genotype frequencies was carried out for variants significant in association analysis. Quality control was carried out on the total cohort of 316 genotyped patients. Samples and variants were excluded if call rates <0.9, minor allele frequency (MAF) <0.01 and/or deviating from Hardy-Weinberg equilibrium (HWE) (P-value <0.0001). The five copy number variants, 46 X-chromosomal variants and one tri-allelic variant present on the array were not included in the analyses.

Isolation of germline DNA in the validation cohort has been previously reported [14,15]; from the Australian validation cohort germline DNA was isolated from blood using the QIaamp DNA Blood Mini kit (Qiagen) according to the manufacturer’s instructions. Genotyping of the validation cohort was performed for six of the ten variants that showed significance in the discovery cohort, excluding four variants that were in linkage disequilibrium (based on r²≥0.8) with any of the six variants and that had higher P-values
than their linked variants. Genotyping of the second validation cohort was subsequently performed for the four variants that were significant in the first validation stage and that showed the same direction of effect in the discovery and validation cohorts. KASP-On-Demand (KOD) assays were used for CYP4F12 rs688755, SLC22A5 rs274548, and FMO6 rs7886938; KASP-By-Design (KBD) assays were used for CYP8B1 rs6771233, SLC22A2 rs316003, and SLC7A8 rs1884545, all according to the manufacturer’s protocol (LGC Genomics, Hoddesdon, UK). Fluorescence was measured with a 7500FAST Real-Time PCR System (ThermoFisher, Nieuwegein, The Netherlands). Genotypes were scored using the algorithm and software (v2.0.6) supplied by ThermoFisher. Blanks (3%) as well as duplicates between plates were included as quality controls for genotyping.

**Statistical analysis**

Statistical differences in demographic data between patients with progressive disease and control patients were assessed by the Fisher exact, Pearson chi-square or Mann-Whitney U tests as appropriate using SPSS v22 (SPSS Inc., Chicago, Ill, USA). To assess the effect of a genetic variant on occurrence of progressive disease, the data was dichotomized to progression yes/no (yes: patients with progressive disease; no: control patients). Associations between genetic variants and the occurrence of progressive disease were assessed by multivariable logistic regression analysis in PLINK using the command --logistic (additive model) (PLINK v1.07) [18]. For the genetic variants that were significantly associated with progressive disease in the discovery cohort, we also assessed potential associations with two other clinical endpoints: recurrent disease (using PLINK), and 5-year disease free survival (DFS) (time interval from diagnosis to either progression or recurrence) using Cox proportional hazards models in SPSS. These variants were excluded from subsequent analysis, to filter out those variants that were not specific for the inadequate drug response observed in progressive patients. Reported P-values are two-sided and are considered statistically significant if <0.05 in the genetic analyses (<0.1 for selection of clinical covariables). No correction for multiple testing was performed because of the exploratory nature of the study. Meta-analysis of the association analysis results of the discovery and validation cohorts, and of all three cohorts including the second validation cohort, was performed using a fixed effects model in PLINK.

To investigate whether the variants associated with progressive disease are also related to 5-year overall survival (OS) (as patients with progressive disease are generally expected to die), we assessed the effect of a combination of genetic variants on the 5-year OS, by calculation of a genetic risk score as previously described [19]. The score is the sum of the number of unfavorable alleles (i.e. alleles inducing risk of progressive disease) carried by each patient for each of the variants significantly associated with progressive disease in the meta-analyses. The influence of the risk score on 5-year OS was visualized using Kaplan-Meier survival curves.

**Results**

**Patient population**

The study was carried out following a three-stage design, including a discovery cohort, an independent validation cohort and an independent second validation cohort. Of the 287 eligible patients in the discovery cohort, four patients were excluded based on a genotype call rate lower than 0.9, leaving 283 patients for analysis. From the validation cohort, all 146 patients were successfully genotyped; from the second validation cohort, one of the 28 patients was excluded based on genotyping failure for all variants, leaving 27 patients for analysis. The patient characteristics of the three cohorts are provided in Table 1. Progressive disease was observed in 13.8%, 12.3% and 18.5% of patients in the discovery cohort, validation cohort, and second validation cohort, respectively.

In the discovery cohort, of all clinical variables included in Table 1, higher age at diagnosis (P=0.058), male gender (P=0.003), the presence of primary metastases (P<0.001), and poor histologic response (P<0.001) were significantly associated with the occurrence of progressive disease and were therefore included as clinical covariables in the genetic analyses, with the exception of the histologic response. As the histologic response and progression are both a reflection of the response to chemotherapy, inclusion of the histologic response as covariable would unintentionally remove variation between progressive and control patients and would therefore result in overcorrection.

In the discovery cohort, 11 (28.2%) of 39 patients with progressive disease survived despite progression; nine of these patients had a follow-up of >5 years. Seven of these patients were classified as progressive because of progression of the primary tumor, which was in all cases followed by interruption of chemotherapy and early invasive surgery. The other four patients showed progression of metastases (N=3), or progression of both of the primary tumor and metastases (N=1); in three of these cases surgical removal of the metastatic lesions has been reported. In the validation cohort, six (all with follow-up of >5 years) of 18 patients survived progression of the primary tumor (N=1), metastases (N=3), local recurrence (N=1) or unknown (N=1). In the second validation cohort, two of five patients survived progression either in the form of local recurrence (follow-up <1 year) or metastases.
Pharmacogenetics of progressive osteosarcoma

After quality control, 710 variants were included in the analysis of the discovery cohort. In the discovery cohort, one of the six genetic variants showed a significant association with progressive disease (Table 3). For this variant, rs1884545 located in the Solute Carrier Family 7 (Amino Acid Transporter Light Chain, L System) Member 8 (SLC7A8) gene, a protective effect was observed for patients carrying the T allele (P=0.018). The effect remained significant with addition of inclusion site (Spain or England) as covariable (P=0.018).

Validation results
In the validation cohort, six variants were investigated based on the results of the discovery cohort. All variants were in HWE (P>0.05) and showed an average call rate of 0.98. Upon multivariable logistic regression analysis of the validation cohort similar to the discovery cohort, one of the six genetic variants showed a significant association with progressive disease (Table 3). For this variant, rs1884545 located in the Solute Carrier Family 7 (Amino Acid Transporter Light Chain, L System) Member 8 (SLC7A8) gene, a protective effect was observed for patients carrying the T allele (P=0.020). The effect remained significant with addition of inclusion site (Spain or England) as covariable (P=0.018).

Meta-analysis
In a meta-analysis of the association analysis results of the six variants genotyped in both the discovery and validation cohort, the independently validated SLC7A8 rs1884545 variant showed a significant protective effect in patients with progressive disease with an odds ratio (OR) of 0.16 (95% confidence interval 0.05-0.50) (Table 3). In both the discovery and validation cohorts, none of the patients with progressive disease were homozygous for the T allele, whereas in the control groups of the discovery and validation cohort, 1.2% and 2.2% of patients were homozygous, respectively. In addition, a substantially lower percentage of patients were heterozygous in the progressive group.
Pharmacogenetics of progressive osteosarcoma

In addition to the SLC7A8 variant, three other variants (CYP8B1 rs6771233, SLC22A2 rs316003, SLC22A5 rs274548) were significantly associated with progressive disease in the meta-analysis, showing a stronger association compared to the results of the discovery or validation cohort alone.

On the basis of the meta-analysis results of the discovery and validation cohort, four variants were investigated in a second validation cohort, including 27 additional patients. In a meta-analysis including all samples investigated, a significant association of SLC7A8 rs1884545 with an OR of 0.22 (0.07-0.63) was observed, corresponding to an approximately four to five-fold protective effect, although the association did not become stronger compared to the meta-analysis without these additional 27 patients (Table 3). The other three variants, SLC22A5 rs274548, CYP8B1 rs6771233, and SLC22A2 rs316003, remained significantly associated with progressive disease, with the latter two showing a stronger association.

After calculation of a genetic risk score based on the four variants associated with progressive disease, a decreasing 5-year OS was observed with an increasing genetic risk score (Figure 1). The 5-year OS curves for each individual variant associated with progressive disease did not reflect a difference in OS between the genotypes (Supplementary Figure S1). However, based on the observation of a combined effect of the four variants, we constructed the genetic risk scores starting with the two variants (rs274548 and rs316003) showing some divergence of the survival curves. The group with the highest genetic risk score combining the effect of the three variants rs274548, rs316003 and rs6771233, revealed the lowest OS, comparable to the highest genetic risk score group combining all four variants, indicating that these three are the variants that contribute to OS.

<table>
<thead>
<tr>
<th>Validation cohort</th>
<th>Meta-analysis</th>
<th>Meta-analysis overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNP</td>
<td>Gene</td>
<td>Chromosome</td>
</tr>
<tr>
<td>rs1884545</td>
<td>SLC7A8</td>
<td>14</td>
</tr>
<tr>
<td>rs6771233</td>
<td>CYP8B1</td>
<td>3</td>
</tr>
<tr>
<td>rs316003</td>
<td>SLC22A2</td>
<td>6</td>
</tr>
<tr>
<td>rs274548</td>
<td>SLC22A5</td>
<td>5</td>
</tr>
<tr>
<td>rs7886938*</td>
<td>FMO6</td>
<td>1</td>
</tr>
<tr>
<td>rs688755*</td>
<td>CYP4F12</td>
<td>19</td>
</tr>
</tbody>
</table>

OR, odds ratio; 95% CI, 95% confidence interval; NA, not analyzed.

Table 3. Association analysis results validation cohort and meta-analyses.
Pharmacogenetics of progressive osteosarcoma

Discussion

Progressive disease occurs in osteosarcoma when there is failure of response to standard chemotherapy. The background of this lack of response is poorly understood. As non-response to chemotherapeutic treatment might be explained by genetic variations in genes encoding metabolizing enzymes and drug transporters, a pharmacogenetic approach could provide important insight in the underlying molecular mechanisms. Therefore, we performed an exploratory study and provide evidence that specific genetic variants involved in drug metabolism and transport are specifically and only associated with progressive disease. We identified four genetic variants associated with progressive disease, of which one variant in \( SLC7A8 \) was confirmed in an independent validation cohort. In addition, three variants in other genes showed the same direction of effect in the discovery and validation cohorts and were significantly associated in meta-analyses.

Until now, a few studies have examined a small number of candidate genes involved in cisplatin pharmacology for association with poor response (defined as stable or progressive disease) in small osteosarcoma cohorts [8-12]. The approach of the present study is slightly different with respect to the grouping of patients, as we focus on the subgroup of patients that do not show an adequate response to standard treatment for osteosarcoma. This subgroup, referred to as patients with progressive disease, includes patients with disease progression under treatment or the lack of complete remission (including partial response or stable disease). This is a slightly broader subgroup than progressive patients according to the RECIST criteria, to include all patients showing an inadequate drug response [20]. To ensure identification of genetic variants specific for this subgroup of patients, we analyzed whether the variants associated with progressive disease could also be linked to the patient group that was initially disease free after treatment, but experienced recurrent disease later on. One can imagine that the same genes or mechanisms, to varying extents, could play a role in both progressive and recurrent disease. For the same reason, we excluded variants significantly associated with 5-year DFS.

The present study is the first large scale study investigating a wide range of genes involved in drug metabolism and transport in relation to progressive disease. The top associated variant is located in the \( SLC7A8 \) gene that encodes an amino acid transporter present in the basolateral membrane of the proximal tubule in the kidney, as well as in the colon and intestine [21]. Ovarian cancer cells resistant to treatment including cisplatin and doxorubicin showed reduced expression of \( SLC7A8 \) [22]. In addition, two other genetic variants in this gene that are not in LD with the rs1884545 variant have been linked to outcome after platinum-based therapy in esophageal cancer [23]. For both the synonymous rs1884545 variant and the linked rs8013620 variant, the functional effect is unknown; it is possible that another linked variant in close proximity is the causal variant.
Notably, two other genes (SLC22A2, Solute carrier family 22 member 2 and SLC22A5, Solute carrier family 22 member 5) of the four identified genes also encode transporters that function in the kidneys. SLC22A2 is implicated in the transport of cisplatin in tubular cells. A genetic variant in the gene (rs316019, not in LD with rs316003) has been linked to cisplatin induced nephrotoxicity [24-26]. SLC22A5 is involved in the reabsorption of carnitine by the proximal tubular cells. Studies have indicated that cisplatin inhibits SLC22A5 functioning also leading to nephrotoxicity [27,28]. Unfortunately, we could not further investigate the relation between the identified variants in SLC22A2 and SLC22A5 and nephrotoxicity, as the availability of information on the adverse drug event in our study cohort was only limited. The fourth identified gene (CYP8B1, Cytochrome P450 8B1) is expressed in hepatocytes and is involved in bile acid production and glucose homeostasis [29]. Thus far, no clear connection between this gene and chemotherapy treatment has been reported. The SLC22A5 rs274548 variant was significantly associated with 5-year progression-free survival in a previous DMET analysis of our group [30]; however, the association was not significant in the present study involving similar analysis (of DFS), probably due to differences including analysis in the combined Dutch cohort rather than meta-analysis of two individual cohorts, and inclusion of multiple clinical covariables in addition to metastases at diagnosis in the present study.

Our data indicate that the link between progressive disease and overall survival is not complete. Although we acknowledge that overall survival is not a pure pharmacogenetic endpoint, especially the top hit SLC7A8 rs1884545 did not contribute to prediction of overall survival, suggesting that progressive disease and overall survival are different disease outcome entities. This is reflected by the observation that not all patients with progressive disease died. However, the survival curves of the genetic risk score of the other three variants suggested a cumulative contribution to overall survival which is not reflected in the survival curves of the individual variants.

We have studied the largest osteosarcoma cohort for pharmacogenetic research regarding progressive disease to date and are the first to include independent validation cohorts. Nevertheless, the patient numbers are still relatively low for genetic association studies and the number of patients with progressive disease is rather low, which is in line with the low incidence of the disease and of disease progression. Therefore, we consider the current study as an important but first step into the pharmacogenetic background of an inadequate drug response in patients with osteosarcoma. As we have retrospectively included patients diagnosed over the past decades, during which imaging techniques have improved, it is possible that we have missed some cases of progressive disease in patients diagnosed in the early years. In addition, the discovery cohort was heterogeneous regarding treatment protocols. A proportion of the patients received two drugs (doxorubicin and cisplatin), whereas other patients also received other drugs in addition to cisplatin and doxorubicin (mostly only MTX), which could give a more favorable outcome. However, there was no significant effect of the presence of MTX in the treatment regimen on progressive disease in our cohort, which makes it likely that the influence of the differences in treatment regimens on the results is limited. Furthermore, because we studied germline variants, we could have missed tumor-specific mutations in genes involved in the uptake of the chemotherapeutic drugs, although studies on genetic variants in genes involved in drug metabolism and transport showed high concordance between DNA derived from tumor and blood or saliva [31]. In addition, the tumor genetic background is also likely to define the intrinsic response to chemotherapy [32].

The goal of identifying pharmacogenetic variants involved in progressive disease is to enable upfront identification of the subgroup of patients showing a very poor response to treatment with conventional chemotherapy. The variants identified in the present study are as single variants not yet discriminative for implementation in the clinical setting, but in the future a combination of these germline variants with tumor intrinsic (genetic) factors could be used to draw a predictive profile for each patient. After upfront identification of patients, other treatment options could be considered which have shown beneficial effects in clinical trials including mostly progressive patients, e.g. 45% 6-months progression free survival with sorafenib and mTOR inhibitor everolimus [7]. Other second-line treatments could also be considered as first-line treatment for this subgroup of patients [2-6]. In addition to such a clinical implementation, these genetic association studies are important to gain more insight into the mechanisms of action of the drugs investigated. This is illustrated by the outcome of the present study in which three of the four genes are implicated in the transport of cisplatin. In addition, these results may provide new leads for development of agents to modulate the response to chemotherapy.

From the patient’s perspective, it will be important to identify the patients at risk of a poor response to chemotherapy. This study illustrates that it is indeed possible to distinguish this patient group from good responders, as we identified pharmacogenetic variants specifically associated with progressive disease in patients with osteosarcoma. Ultimately, clinical implementation of a validated (genetic) risk profile may enable treatment strategies specifically targeting this subgroup of patients.
References


Supplementary data

Figure S1. Kaplan-Meier survival curves for genotypes associated with progressive disease. Curves of the 5-year overall survival are depicted for each of the four genetic variants significantly associated with progressive disease in meta-analysis of the discovery and validation cohorts.
Chapter 8
General discussion
Scope and aim of thesis

The treatment of osteosarcoma poses a major burden to patients because of the combination of often disabling surgery and toxic chemotherapy with the risk of serious long-term side effects. At diagnosis, the patient’s response to chemotherapy with regard to efficacy and toxicity, is largely unknown and is predominantly a matter of close monitoring during and after therapy, as large variation exists between patients’ drug responses. Drug response is a complex trait and the increasing knowledge of genetics has led to the understanding that genetic factors also contribute to variation in drug response. In cancer, both the somatic and germline genome are relevant for treatment outcome [1]. Whereas the somatic genome has been a major focus area in research to understand therapy resistance and efficacy, the germline genome can also be highly relevant for treatment efficacy and toxicity through involvement in drug exposure and drug effects. The science of pharmacogenetics aims to facilitate personalized medicine by concentrating on the influence of heritable genetic variation on drug response, with the purpose to predict drug efficacy or toxicity at diagnosis using genetic biomarkers. In addition, it can increase the understanding of the interindividual variation in drug response, and provide new therapeytic targets. The aim of this thesis was to gain insight into the germline genetic factors underlying chemotherapy efficacy and toxicity in patients with osteosarcoma. This could ultimately contribute to more effective treatment and prevention of undesired side effects. We have focused on validation of genetic associations reported in literature by ourselves and others as well as on discovery of novel associations. In this final chapter, the findings of the research presented in this thesis are discussed, followed by a discussion of the studies’ strengths and limitations, and a future perspective on further research and the clinical setting.

Main findings

For pharmacogenetic analysis of Dutch osteosarcoma patients as presented in this thesis, we retrospectively recruited patients diagnosed from as early as 1975 up to 2013. As a proportion of these patients was not alive at the time of inclusion, from these patients the only available source of germline DNA was normal formalin-fixed, paraffin-embedded (FFPE) tissue that had been archived for many years. In chapter 2, we have validated the use of germline DNA isolated from normal FFPE tissue for genotyping with the Drug Metabolizing Enzymes and Transporters (DMET) Plus array, by comparing genotypes obtained with blood or saliva-derived DNA with those of normal FFPE tissue-derived DNA from sixteen alive osteosarcoma patients. We have been the first to report that DNA isolated from archived normal FFPE tissue samples can be reliably used as a template for the DMET Plus array. This finding was an essential technical validation step for our subsequent pharmacogenetic studies using this platform. Moreover, it adds to the existing literature on the suitability for the use of DNA from FFPE tissue samples with a variety of genotyping technologies [2-5], which unlocks the opportunity to access genetic information of large retrospective series of samples for use in pharmacogenetic studies. This is especially valuable for establishing unbiased and well-powered cohorts in retrospective studies of cancers with low incidence and survival, as prospective sample collection within clinical trials, which has become more routine, will take years to establish large cohorts of rare diseases such as osteosarcoma.

In the study reported in chapter 3, we have investigated a variant in the ACYP2 gene, which was identified as a pharmacogenetic marker of cisplatin-induced ototoxicity in patients with brain tumors in a recent genome-wide association study (GWAS) by Xu et al. Cisplatin-induced ototoxicity is a serious clinical problem with large interpatient variability, which cannot be fully explained by clinical or genetic risk factors identified thus far. We have been the first to replicate the finding of Xu et al. in an independent cohort of 156 cisplatin-treated patients with osteosarcoma. In line with the previous publication, all patients carrying the A allele of the genetic variant in the ACYP2 gene presented with ototoxicity. We have not only confirmed the association itself but we have also shown that it is generalizable among other cancers. Therefore, the genetic variant is of clinical relevance and provides essential evidence for the genetic basis of cisplatin-induced ototoxicity, which has proven to be difficult to pinpoint in previous pharmacogenetic studies. Although the ACYP2 variant should be considered a useful marker for upfront identification of patients susceptible to ototoxicity based on our results, the frequency of the variant is low. Hence, there is a proportion of patients presenting with ototoxicity without having the ACYP2 variant, and thus the variant covers only a part of the variation observed in the susceptibility to ototoxicity.

Following a previous candidate gene analysis of our group investigating methotrexate (MTX) plasma levels and toxicity in a small cohort of osteosarcoma and acute lymphoblastic leukemia patients (thesis Melanie Hagleitner [6]), we have further explored the role of absorption, distribution, metabolism, and excretion (ADME) genes in MTX-induced toxicity using the DMET Plus array in exclusively osteosarcoma patients, as presented in chapter 4. We have focused specifically on hepatotoxicity because this can be related to MTX but not to the other agents of the osteosarcoma treatment regimen. None of the 32 genetic variants that were associated with high 48 hour MTX plasma levels, were also associated with MTX-induced hepatotoxicity. However, when we focused on the extreme cases, i.e. patients with frequent hepatotoxicity (elevated liver enzymes after ≥4 of 12 MTX courses), we found a protective effect of the GSTM3 rs1799735 (delAGG) variant (GSTM3*1B) for frequent hepatotoxicity. GSTM3 codes for a glutathione S-transferase, an enzyme involved in detoxifying electrophilic xenobiotics, including antitumor agents,
by conjugation with glutathione [7]. The GSTM3*1B allele, which is suggested to lead to increased levels of the GSTM3 enzyme, has not been investigated in previous studies into the pharmacogenetics of MTX plasma levels and hepatotoxicity in osteosarcoma [8-14]. Interestingly, the same variant has also been reported to protect from cisplatin-induced ototoxicity in osteosarcoma [15], indicating that the effect of the variant may be favorable in the context of toxicities of different drugs included in the treatment protocols for osteosarcoma.

In chapter 5, we have reported on the confirmation of the role of a genetic risk score that we have previously reported, consisting of several pharmacogenetic markers associated with the treatment efficacy in osteosarcoma patients [16]. Even in the absence of reliable genotypes for the CYP3A4 variant, which was one of the five previously identified markers, we were able to reproduce the association of a genetic risk score based on genetic markers in Fasl, MSH2, CASP3, and ABCC5 with 5-year progression-free survival (PFS) in independent osteosarcoma cohorts. Additionally, in a meta-analysis also including the two cohorts of the original study, we confirmed the association of two of the individual variants (in MSH2 and ABCC5) with 5-year PFS. Although the validation was not complete, in the sense that the association of the genetic risk score with outcome was not statistically significant in the individual cohorts but only after combining cohorts, and that not all individual variants were significant in the meta-analysis, the findings highlight that even in relatively small patient cohorts it is possible to identify clinically relevant markers. When compared to other pharmacogenetic studies into drug efficacy in osteosarcoma, this is a unique finding, even though further validation in larger datasets is needed.

The DMET analysis reported in chapter 6 represented a next step in the search for the pharmacogenetic basis of drug efficacy in osteosarcoma. Eight genetic variants were identified in relation to 5-year PFS, from which a variant in the ABCB1 gene was independently validated in one of the validation cohorts. Seven other variants were significantly associated with 5-year PFS in a combined meta-analysis of discovery and validation cohorts. Our findings are an addition to other reports indicating that ABCB1 contribute to differential drug efficacy in osteosarcoma patients [17-21]. It is also the first indication that variation in genes of the solute-linked carrier family are involved in drug efficacy in osteosarcoma. Furthermore, none of the variants that were associated with histologic response (HR) in our two Dutch cohorts, could be validated in independent cohorts. Also in the previous pathway-based study of our group, no variants relevant to the HR could be identified. There could be several explanations for this. One potential explanation is that it may be possible that the genetic basis of HR should be searched for mainly in the somatic genome, which largely dictates the tumor behavior and response. It could also be possible that pharmacodynamic genes (at germline level) rather than pharmacokinetic genes as studied in the DMET analysis and other than the pharmacodynamics genes studied in the previous pathway-based approach are involved in the variability in histologic response. In addition, the HR may not be a suitable endpoint for (retrospective) analysis for example because of heterogeneity due to sensitivity to interobserver differences.

In chapter 7, we have reported on the results of a DMET analysis as the first step to identify patients for whom standard chemotherapeutic treatment is ineffective, of which the background is poorly understood. We have focused on the subgroup of patients showing progressive disease during or shortly after treatment and identified genetic variants involved in drug metabolism and transport that were specifically and only associated with progressive disease. Of four genetic variants associated with progressive disease, an SLC7A8 variant with a protective effect was confirmed in an independent validation cohort. In addition, three variants in other genes (CYP8B1, SLC22A2, SLC22A5) were significantly associated in a meta-analysis. The SLC7A8 transporter has been linked to cisplatin and doxorubicin resistance and outcome after platinum-based therapy in other cancers, and it is therefore considered a plausible outcome of this exploratory study deserving further follow-up [22,23].

In the present thesis, we have described studies using the DMET array focusing on either toxicity or efficacy endpoints. Of note, there is some overlap regarding the involvement of genes in MTX plasma levels (chapter 4) and in the meta-analysis results of 5-year PFS analysis (chapter 6), namely ABC4 and SLC22A14. In the MTX-treated patient population studied in chapter 4, both genes were associated with MTX plasma levels (though not with hepatotoxicity). However, there was no association of MTX plasma levels (which is influenced by pharmacokinetics) with 5-year PFS (which might be influenced by pharmacokinetics, pharmacodynamics, or both), making it difficult to interpret the relevance of these genes in the context of MTX treatment. The association of these genes with 5-year PFS might be related to transport of other chemotherapeutics used in the treatment of osteosarcoma, especially for ABC4 for which a link to cisplatin-resistance has been reported in other forms of cancer [24].

**Strengths and limitations**

Some of the most important determinants of an informative outcome of observational research in the medical sciences are sufficient sample size and representative sampling of the patient population under study. Genetic association studies integrate clinical data from medical files and genetic information of each patient; of which the latter is the most critical and challenging factor for establishing an informative patient cohort. In the context of osteosarcoma research, this is even more challenging because total patient numbers are limited by the rarity of the malignancy. In addition, a proportion of patients
is unavailable for sampling in a retrospective setting because of the high mortality of the disease. From this perspective, large numbers of osteosarcoma patients were included in the studies presented in this thesis, with up to 310 Dutch patients the largest cohort used in pharmacogenetic research in osteosarcoma in the world, and inclusion of study cohorts of also exclusively osteosarcoma patients from collaborative research groups. Moreover, to maximally prevent bias, every effort was made to include the (Dutch) patients who had died before start of the studies, by collecting samples from the pathology archives. Nonetheless, patient numbers were still relatively low for genetic association analysis, which could result in false-negative outcomes by missing true associations, or underestimation of the strength of associations, for example of the genetic variants that were significantly associated with PFS or progression in meta-analysis of multiple cohorts, but could not be reproduced in individual validation cohorts. On the other hand, the findings of mainly the ACYP2 study (chapter 3) and DMET analysis of PFS with \textit{ABCB1} as outcome (chapter 6), nicely illustrate that the investigated cohorts were of sufficient sample size to detect associations of genetic variants with low frequency or small effect size. The limited number of patients, in combination with the low frequency of the phenotype under study, as was the case for patients progressive disease (chapter 7), has determined the design of that particular study, which included combining multiple cohorts into a single cohort for statistical analysis, making this an exploratory study.

A rational choice of genes to study, which may include selection of meaningful biological pathways or instead taking a genome-wide approach, depending on the existing knowledge, patient numbers, and resources, is another determinant of the value of a genetic association study. The main research described in this thesis included analysis with the DMET Plus array. The DMET gene selection was based on a core list of ADME genes defined by academic, industry, and genomic experts in the PharmaADME consortium, with addition of genes regulating intracellular processes that facilitate ADME [25]. The variants within these genes were ranked according to clinical research utility. Following small candidate gene studies reported in literature and a candidate pathway-gene based approach previously performed by our group, with all of these studies selecting genes for their known function in the processing of the chemotherapeutics used in osteosarcoma treatment, the studies presented in this thesis involving DMET analysis are the first to include such a substantial number of genes. Although this is still a form of a candidate gene selection since all genes are related to drug absorption, metabolism, excretion and transport, the array also includes metabolic or transporter genes without a direct link with the chemotherapeutics involved, which enables discovery of novel associations. Related to that is the fact that several of the genetic variants previously described in osteosarcoma pharmacogenetics literature (such as several \textit{ABCB1} and \textit{GSTP1} variants related to drug efficacy) were analyzed in our cohorts because they are present on the DMET Plus array. Hence, although our DMET analyses were designed to discover novel associations, it has also given the opportunity to assess previously described associations, especially as the size of our Dutch cohort(s) was suitable for validation of markers described in literature, which is illustrated by the ACYP2 study. The downside of this platform is that it includes a large number of rare variants, which are included because they are presumably more likely to be causal than common variants. However, they were monomorphic or had a frequency too low to include in association analysis, because extremely low minor allele frequencies would require substantially larger patient cohorts to have sufficiently powered analyses. This was not only the case in our studies but also in other association studies using this platform [23,26-28]. This substantially decreased the number of genetic variants eligible for analysis. Furthermore, the selection of genes included in the array is based on the gene function in pharmacokinetics and is therefore not completely hypothesis-free [25]. Therefore, we might still have missed associations of pharmacodynamic-related genes or of genes affecting drug response through other mechanisms.

When interpreting the results of genetic association studies, investigators must always be aware of the risk of false-positive findings, especially when large numbers of genetic variants are tested. Validation in independent cohorts is therefore an essential component to overcome the winner’s curse and be able to appreciate the results and strengthen the evidence. In the studies related to drug efficacy presented in this thesis, we have included a validation phase with patient cohorts of collaborating research groups from outside of The Netherlands. We acknowledge that the cohorts were not ideal validation cohorts with respect to size, sampling, and differences in treatment regimen (i.e. different agents in addition to the standard regimen of cisplatin and doxorubicin). However, the papers presented in this thesis are the only studies within the osteosarcoma pharmacogenetics field reported so far that include (multiple) validation sets at all. Moreover, as any osteosarcoma cohort is small because of the rarity of the disease, the validation steps as presented here were the best available options to date. Small validation cohorts carry the danger of false-negative findings leading to incorrect conclusions by missing true associations. It is therefore far from obvious that we were able to confirm genetic associations in these cohorts, although it is still possible that other true associations have been missed. Another main issue encountered in the validation analyses was the difference between cohorts with respect to the phenotype under study, with the English cohort showing a remarkably high PFS, which was most likely predominantly reflecting a selection bias towards survivors. With hindsight, we realized that this cohort might not have been a suitable choice for validation of markers associated with PFS. Unfortunately, the validation cohorts used in the drug efficacy studies could not be used for our study into MTX-induced hepatotoxicity, as no information was available on the liver enzymes of these patients.
The genetic association studies presented in this thesis were all case-control studies, a widespread used study design in the field of genetic epidemiology, in which patients are classified based on their clinical drug response (i.e. efficacy or toxicity), and in which cases and controls are expected to differ in their prevalence of genetic variants related to drug effects. As cases and controls might also differ in other (clinical) factors, such as the presence of metastases at diagnosis which is also related to PFS, the endpoint used to analyze drug efficacy, we have adjusted for these factors where needed, to remove variation between cases and controls that is not directly related to the genetic makeup. Nevertheless, a potential difficulty is the possibility that other yet unrecognized clinical or biological factors represent subgroups of patients or tumors, that were not accounted for in the analyses, and which could dilute or obscure true associations. In addition, we have focused on the germline genome, not taking into account the tumor genetics, which could be of influence on the efficacy endpoints (histologic response, PFS) by determining drug resistance or natural tumor behavior. However, it is clear from the current literature that the complex genomic profile and genomic instability of the tumor are highly challenging in research aiming to define correlations of genetic tumor profiles and survival. To date, there is no distinct, reproducible genetic profile reported of osteosarcoma which could be screened for and included in our studies. Moreover, it is exactly because of the large heterogeneity within and between tumors, that it is unlikely that a substantial proportion of patients in our cohorts would harbor the same yet undefined poor prognostic genetic tumor profile. Hence, the influence of the tumor genetics is not expected to have largely influenced our results. Furthermore, although concordance between germline and tumor DNA has been described for other tumors [29], it could be a relevant step to validate the findings from germline studies into the treatment efficacy in osteosarcoma patients using tumor DNA, depending on the gene function, which has not been carried out in the described studies.

An intrinsic limitation of genetic association studies is the fact that the results do not provide an answer to the question as to whether an associated variant itself is of causal influence on the phenotype under study, or whether another variant in linkage disequilibrium with the associated variant is so, although in case of associated variants with a known functional effect, the former becomes more likely. Additionally, in the context of multidrug treatment, it also remains a question as to which drug the genetic effect is related, or whether it is related to multiple drugs, which generally remains speculation until mechanistic evidence supporting the findings is provided. Furthermore, with the potential future use of pharmacogenetic variants as screening tool in mind as an ultimate aim of pharmacogenetic analyses, the causality of a specific drug in adverse drug reactions is also a point of interest. It is for this reason that specifically hepatotoxicity induced by MTX was studied (chapter 4), and other toxicities such as bone marrow depression, which is also a common side effect of the other agents in the MAP (MTX, doxorubicin, cisplatin) treatment regimen, were not studied. Furthermore, based on the indications that MTX plasma levels are related to MTX-induced toxicities, and also to narrow down the range of genetic variants included in genetic association analysis with hepatotoxicity, we first analyzed the association of genetic variants with high MTX plasma levels (chapter 4). As it became clear from our cohort that MTX-induced hepatotoxicity not necessarily presents in patients with high MTX plasma levels, the downside of the approach is that we might have missed genetic variants related to MTX-induced hepatotoxicity that is not a consequence of high exposure. In fact, this brings up new questions from a pharmacological point of view about the mechanisms of MTX-induced hepatotoxicity, as it seems surprising that it is not (necessarily) a result of high MTX plasma levels. With passive diffusion being dependent on the extracellular MTX concentration [30,31], could it be that in patients with rapid clearance, hence with relatively low MTX plasma levels, there is relatively higher active MTX uptake in the liver, contributing to hepatotoxicity? Furthermore, as MTX is subject to intracellular polyglutamation in liver cells, which prevents it from moving out of the cell, the extent of polyglutamation of MTX might also be of influence [32].

Last of all, we analyzed only those patients with available clinical data, ignoring missing data. Although survival data were available for virtually all included patients, data on adverse events (ototoxicity and hepatotoxicity) was lacking for a proportion of patients, even though the large majority of patients had been treated within a clinical trial with close monitoring. In general, the availability of toxicity data might be biased towards patients experiencing toxicity leading to increased monitoring, although in our study cohort we did not observe an extremely high frequency of ototoxicity in the included patients. Furthermore, pharmacogenetic analysis of doxorubicin-induced cardiotoxicity was not carried out because of the absence of uniform measurements of a substantial number of patients.

Future perspective

Research perspective

The pharmacogenetic studies in osteosarcoma reported in literature so far give some clues on the genetic component of variability in response to chemotherapy in the treatment of osteosarcoma, but also illustrate the complexity of pinpointing the most relevant genes, especially in the context of a complex phenotype as drug response and in a rare disease. The research presented in this thesis has provided several validated genetic markers of chemotherapy efficacy or toxicity, indicating that a pharmacogenetic approach is able to yield reproducible insights in osteosarcoma therapy response. Nevertheless, to some extent the DMET analyses did not yield truly unexpected findings with respect to...
the validated markers, for example ABCB1 variants in relation to therapy outcome have been studied before as a candidate gene. This raises the question as to whether the genetic effects that have been identified until now, in fact represent all the genetics underlying the variability in drug response, or alternatively, whether pharmacokinetics, which was the focus of the DMET studies, is just not the mechanism in which to search further. In addition, any further explanation of variability in drug response might also need to be searched for in non-genetic factors, for example environmental factors such as nutrition, tobacco smoking, alcohol use, co-administered drugs, and environmental chemicals. In case the current findings represent all pharmacokinetics related genetic effects in osteosarcoma treatment, it would be worthwhile to more extensively study pharmacodynamic genes or to go beyond typical pharmacokinetic and pharmacodynamic pathways and more forward to scanning the whole genome by a GWAS or next generation sequencing. Such truly hypothesis-free methods have the potential to discover novel associations of unexpected genes with drug response. Whereas genome-wide approaches have been exhaustively carried out in many other diseases, often without clinically relevant and applicable results, this type of analysis is yet to start in osteosarcoma pharmacogenetics. Although a GWAS has been carried out in osteosarcoma to detect disease or metastasis susceptibility genes, no pharmacogenetic endpoints were included, presumably because of lacking clinical drug response datasets [33,34]. As the first GWAS into cisplatin-induced ototoxicity in a relatively small population of pediatric patients has resulted in the identification of a genetic variant (in ACYP2) that could be confirmed in osteosarcoma as described in this thesis, it seems worthwhile to explore this approach further in osteosarcoma. Moreover, pharmacogenetic findings could not only lead to a genetic profile for prediction of drug response, but also holds the opportunity to increase our understanding of drug response in patients with osteosarcoma, and to provide new druggable targets for developing novel agents for patients prone to have a poor response to conventional treatment.

Real progress in the osteosarcoma pharmacogenetics field will require large studies with respect to patient numbers. Mixed cancer populations could provide a good option for large scale studies to detect associations that are not tumor-specific. This is primarily feasible for toxicity studies because drug toxicity is similar among different cancers treated with the same chemotherapeutics, and the effect sizes are expected to be larger than observed in efficacy studies. Mixed cancer cohorts are less suitable for efficacy studies, because efficacy endpoints are difficult to compare between different tumors. Other than combining cancers, the patient population of the recent EURAMOS-1 trial with over 2,200 similarly treated osteosarcoma patients provides an excellent opportunity for large scale pharmacogenetic investigations, as well as for further validation of genetic markers reported in literature and identified in the research described in this thesis [35]. Aside from validation in osteosarcoma cohorts, it could also be worthwhile to translate the findings of variants such as the ACYP2 variant to other cancers with the same treatment, for example adult cisplatin-treated patients with head and neck cancer.

Functional characterization of the effect of genetic variants found in genetic association studies in osteosarcoma, including the research presented in this thesis, is (mostly) lacking. Understanding of the biological mechanism underlying the findings and proving causality will be valuable to provide additional evidence of the genetic findings, especially in the context of the limited sample sizes of osteosarcoma studies. Moreover, it could also shed light on the specific chemotherapeutic agent to which the genetic variant is linked, which may not always be clear in the context of multidrug treatment. In functional studies, a distinction should be made between the effect of genetic variants in the germline or specifically in the tumor. Whereas for liver enzymes the effect of a germline genetic variant can be expected to be restricted to its function in the liver, for drug transporters present in typical drug processing organs such as liver and kidneys but also in the tumor, prediction of the effect of a variant also involves studying the effect at the tumor level using osteosarcoma cell models.

The interpretation of pharmacogenetic data is still a challenge for translation to the clinic. Until now, the efforts in pharmacogenetic studies in osteosarcoma in the recent years have not led to a change in clinical practice. All discovered variants are still in the research phase and warrant further validation. Hence, future studies should be focused not only on independent validation of current findings and discovery of yet unknown associations, but also on the pathway towards clinical implementation, which includes testing of clinical validity and translation through prospective studies demonstrating improvement of outcome [36].

**Clinical perspective**

As survival of osteosarcoma patients treated with conventional chemotherapy has stagnated over the past decades, many efforts are made to identify and develop alternative treatment options. Apart from a variety of ongoing trials [37], there have been a few publications of clinical investigations of new drugs for the treatment of osteosarcoma during the past four years while the research presented in this thesis was conducted. The bisphosphonate zoledronic acid, which targets the bone microenvironment, is in an early trial phase [38], but a first report on clinical experience with zoledronic acid only has shown encouraging results with respect to PFS in metastatic patients [39]. Another type of investigational agent is muramyl tripeptide-phosphatidyl ethanolamine encapsulated in liposomes (L-MTP-PE), which is an immunomodulating agent that has shown positive results with improvement in overall survival in localized and metastatic osteosarcoma [40]. The EURAMOS-1 trial has investigated the immunomodulator interferon alpha 2b as maintenance therapy for good responders to neoadjuvant chemotherapy. The first results do not support routine use of interferon alpha 2b as event-free survival was not
improved in the experimental arm [41]. Furthermore, the use of monoclonal antibodies, which receive much attention in other cancers, is also being tested in the treatment of osteosarcoma. The HER2 targeting monoclonal antibody trastuzumab has not resulted in improved outcome of patients with a HER2 overexpressing tumor in a phase II trial [42]. Another phase II trial has shown positive results of the combination of mTor inhibitor everolimus and angiogenesis inhibitor sorafenib [43]. It is clear that there are only a few promising new agents for osteosarcoma treatment, and it will take time to bring these to clinical practice. Therefore, another parallel approach for clinical research in osteosarcoma is the personalization of treatment with chemotherapy by optimizing the use of the drugs currently available.

How could pharmacogenetic variants be used to guide therapy for osteosarcoma patients? Generally speaking, for patients harboring a genetic risk profile for developing toxicities, increased surveillance and use of protective agents without hampering the antitumor efficiency of the chemotherapeutic drug might be future clinical applications. Dose reductions could be considered, as often variants increasing susceptibility to toxicity might also increase efficacy; however, this requires large confidence of the genetic effect. Specifically, considerations for preventing cisplatin-ootoxicity include using alternative chemotherapeutics such as carboplatin, decreasing cisplatin dose [44], increasing monitoring, or using otoprotective agents such as sodium thiosulfate [45]. The ACYP2 variant reported in chapter 3 as predictor of cisplatin-induced ototoxicity represents a good candidate to be considered for investigation in the setting of future international osteosarcoma trials. As the frequency of the variant is rather low, a suitable approach could be to focus on the subset of patients carrying the genetic variant and extrapolate the results to the patient population as a whole. It is also questionable whether genotyping for this variant could be cost-effective; however, this might be considered incidental in the vulnerable pediatric oncology patient population. In addition, despite the low frequency of the ACYP2 variant, cost-effectiveness might be reachable when considering the high health care costs and high impact on quality of life for patients experiencing (severe) ototoxicity. Clinical application of genetic markers of MTX-induced hepatotoxicity could involve dose reduction to prevent hepatotoxicity; however, this is not without the concern of reducing efficacy as well. In the oncological setting, quite a level of toxicity is accepted in favor of maximizing efficacy. Hence, if MTX dose reductions compromise its efficacy, such a trade-off is certainly undesired and unlikely to be accepted for clinical practice. Moreover, although the genetic variant identified in our study was related to both MTX plasma levels and hepatotoxicity, there was also a major proportion of patients presenting with hepatotoxicity in the absence of high MTX plasma levels, which in fact makes it questionable whether dose reduction would diminish the risk of hepatotoxicity. At the same time, for patients with genetic variants predicting poor efficiency of the treatment, the option of intensification of therapy is a difficult one because of the risk of toxicities. Moreover, to date the intensification by adding alternative agents such as in the MAPIE regimen of EURAMOS-1 has not shown the desired improvement of survival [46]. Instead, a poor predictive profile might help to identify those patients for whom alternative treatment, once it is available, should be favored over the standard (chemotherapy) treatment. In addition, pharmacogenetic profiling might enable the identification of those patients who are now in fact undertreated and could be subjected to more intense treatment. Furthermore, for patients not responding to standard chemotherapy, such as focused on in chapter 7, the question arises as to whether chemotherapeutic treatment should be started at all, or whether prevention of treatment that is ineffective and unnecessarily harmful with respect to side effects should be favored for the patient’s quality of life. For this group of patients, the use of experimental drugs could also be considered. For the small group of patients that survived despite their inadequate drug response in our study cohort, who probably belong to the small proportion of patients that is known to survive due to successful surgery, the choice to restrain from chemotherapy might even be a favorable option with positive outcome. Obviously, this could only be considered in light of large confidence of the predictive capacity of the (genetic) profile, which is truly a future perspective.

Concluding remarks

The studies presented in this thesis have provided interesting results further supporting a pharmacogenetic approach in osteosarcoma research. Only through joint efforts with intensive international collaboration and integration of results from pharmacogenetic research with potential future insights at the somatic level, it might eventually be possible to implement a priori genetic screening for guidance of therapy for the benefit of osteosarcoma patients.
References


Summary
Samenvatting
Summary

Osteosarcoma is the most common bone tumor in children and adolescents. Since the seventies in the last century, patients with this malignant tumor are treated with a combination of surgery and chemotherapy, the latter consisting of at least cisplatin and doxorubicin. Although the addition of chemotherapy to the treatment regimen has had a strong impact on survival, for still around 40% of patients this treatment is inadequate and they do not survive. In addition, survivors frequently experience long-term side effects as a result of the toxic chemotherapeutic treatment. More insight into the genetic factors underlying therapy resistance and/or sensitivity to toxicities could be helpful for upfront risk stratification, optimization of the treatment, and development of new therapeutic approaches. In general, many genes are known to play a role in the interaction of drugs and the body (‘pharmacogenetics’). Subtle variations in these genes can have an influence on this interaction. In the studies presented in this thesis, a potential link between variations in genes related to drug absorption, distribution, metabolism, and excretion (ADME) and the effect of chemotherapy in osteosarcoma patients was investigated with regard to efficacy and toxicity.

Chapter 1 provides a general introduction including a comprehensive overview of published pharmacogenetic research involving patients with osteosarcoma.

Chapter 2 describes a technical validation of the use of DNA isolated from long-term stored tissue from the pathology archives on a specific genotyping technology, the Drug Metabolizing Enzymes and Transporters (DMET) Plus array, which is used in the majority of pharmacogenetic studies presented in this thesis. A proportion of osteosarcoma patients included in our studies was not alive at the time of retrospective sample collection, and therefore the only available source for germline DNA of these patients was normal formalin-fixed, paraffin-embedded (FFPE) tissue. As the long-term storage and tissue fixation procedures can negatively influence the quality of FFPE-derived DNA, we compared the DMET array genotyping results of DNA isolated from archived normal FFPE tissue with DNA from matched blood or saliva from sixteen patients with osteosarcoma. We observed high call rates for the FFPE-derived DNA samples (98.9%) as well as high concordance among the blood/saliva-FFPE DNA pairs (97.4%). We therefore concluded that DNA isolated from archived normal FFPE tissue samples can be reliably used as a template for the DMET Plus array.

The next two chapters describe studies into the pharmacogenetics of chemotherapy-induced toxicities in patients with osteosarcoma. In Chapter 3 we focused on cisplatin-induced hearing loss (ototoxicity), which is an irreversible side effect with large, partly unexplained interpatient variability. In a recent genome-wide association study, a genetic variant (rs1872328) in the ACYP2 gene was identified as pharmacogenetic marker of cisplatin-induced ototoxicity in patients with brain tumors. We confirmed this finding in an independent cohort of 156 osteosarcoma patients by demonstrating a significant association between carriership of the A allele of the variant in the ACYP2 gene and cisplatin-induced ototoxicity. This also shows that the association applies to different types of cancer with different treatment protocols. We concluded that the ACYP2 variant should be considered a predictive pharmacogenetic marker for hearing loss, which may be used to guide therapies for cisplatin-treated patients.

Chapter 4 describes an exploratory pharmacogenetic study focused on high-dose methotrexate (HD-MTX)-induced hepatotoxicity. MTX is a chemotherapeutic frequently used in the treatment of pediatric osteosarcoma, and was one of the standard agents used in the recent EURAMOS-1 trial in osteosarcoma. Delayed clearance of MTX may increase the risk of MTX-induced side effects such as hepatotoxicity. We explored the influence of 1936 genetic variants in 231 ADME genes on interindividual variation in MTX plasma levels and hepatotoxicity in 144 HD-MTX-treated osteosarcoma patients using the DMET Plus array. When we analyzed a total of 1250 HD-MTX courses, we detected associations of 32 genetic variants with 48 hour MTX plasma levels. This included a variety of genes unknown to play a role in MTX pharmacokinetics. Of the genetic variants associated with MTX plasma levels, we identified a protective effect of GSTM3*1799735 (delAGG) (GSTM3*B) in relation to frequent hepatotoxicity. These findings indicate a role of pharmacogenetic variants in the interindividual variation in MTX plasma levels and HD-MTX-induced hepatotoxicity and contribute to pinpointing the specific genes involved. Upon validation of the results, the identified GSTM3 marker might eventually be useful for optimization of MTX treatment.

The next chapters concentrate on the pharmacogenetics of chemotherapy efficacy. In chapter 5 we report on a study assessing the reproducibility of our previously identified association of a genetic risk score, based on several pharmacogenetic markers, with 5-year progression-free survival (PFS) using independent Dutch, Spanish, and English osteosarcoma cohorts. Although one of the variants, located in CYP3A4, could not be reliably determined, we could still reproduce the association of the genetic risk score with 5-year PFS based on markers in Fasl, MSH2, CASP3, and ABCC5. Patients carrying no or one risk allele showed a 5-year PFS of 78.5%, compared to a 5-year PFS of 59.5% in patients carrying 2-3 risk alleles, 48.0% with 4–5 alleles, and 57.1% with more than 5 risk alleles (P=0.040). In addition, we also confirmed the association of two of the individual variants (in MSH2 and ABCC5) with 5-year PFS in a meta-analysis also including the two cohorts of the original study. These pharmacogenetic findings are potentially very useful...
for risk stratification at diagnosis and provide a basis for pharmacogenetic profiling in the future.

As a next step, we performed a DMET analysis to identify (additional) associations of variants in ADME genes with 5-year PFS, described in chapter 6. Upon discovery in 283 Dutch osteosarcoma patients and validation in Spanish and English cohorts, we identified eight genetic variants in relation with 5-year PFS. Of these, the variant rs17064 in the \textit{ABCB1} gene was independently validated. Seven other variants of five drug transporter genes (\textit{ABCC4, ABCC5, ABCG1, SLC22A5, and SLC22A14}) were significantly associated with 5-year PFS in a combined meta-analysis of discovery and validation cohorts. These findings contribute to the indications of other studies reporting the influence of polymorphisms in \textit{ABCB1} on outcome after chemotherapy-based treatment in osteosarcoma, and for the first time indicate a role of solute-linked carriers. These markers are of potential interest for the development of new treatment strategies and optimizing current therapy.

In chapter 7 we describe another DMET analysis in which we focused on the subgroup of patients showing an inadequate drug response, defined as progressive disease during or shortly after treatment. Upon discovery in our Dutch cohort of 283 osteosarcoma patients and validation in Spanish, English, and Australian patients, we identified an independently validated \textit{SLC7A8} variant (rs1884545) specifically and only associated with progressive disease, showing a protective effect. Additionally, three variants in other genes (\textit{CYP8B1, SLC22A2, SLC22A5}) were significantly associated in a meta-analysis. This exploratory analysis represents the first step to identify patients for whom chemotherapeutic treatment is ineffective and should be further explored in additional cohorts.

Chapter 8 provides a summarizing discussion of the research described in this thesis, including a future perspective on further (pharmacogenetic) research in osteosarcoma as well as on the clinical aspects of osteosarcoma treatment.

---

**Samenvatting**

Osteosarcomen zijn de meest voorkomende bot tumoren bij kinderen en adolescenten. Sinds de jaren zeventig van de vorige eeuw worden patiënten met deze kwaadaardige tumoren behandeld door middel van een operatie en chemotherapie, bestaande uit een combinatie gebaseerd op cisplatin en doxorubicine. Hoewel de kans om de ziekte te overleven sterk is verbeterd sinds de introductie van chemotherapie in de behandeling, schiet de behandeling voor een deel van de patiënten tekort waardoor nog ongeveer 40% overlijdt. Daarnaast hebben patiënten die de ziekte wel overleven ook op lange termijn vaak te maken met schadelijke bijwerkingen van de chemotherapie, zoals hartschade, nierschade en gehoorschade. Meer inzicht in de genetische factoren die een rol spelen bij (on)gevoeligheid voor het antitumor effect van de chemotherapie of juist gevoeligheid voor de bijwerkingen, zou kunnen leiden tot een effectievere en veiligere behandeling. In algemene zin is van allerlei genen al bekend dat ze een rol spelen bij de werking van medicijnen (‘farmacogenetica’). Kleine variaties in deze genen kunnen een effect hebben op deze werking. In het in dit proefschrift beschreven onderzoek is in osteosarcoompatiënten gekeken naar genen betrokken bij absorptie, distributie, metabolisme en excretie (ADME) van medicijnen en of bepaalde variaties hierin een verband hebben met het effect van de chemotherapie, zowel met betrekking tot de effectiviteit als de toxiciteit.

**Hoofdstuk 1** is een algemene introductie waarin een uitgebreid overzicht wordt gegeven van de bestaande literatuur met betrekking tot farmacogenetisch onderzoek in osteosarcompatiënten.

In **hoofdstuk 2** wordt een technische validatie beschreven voor het gebruik van DNA, geïsoleerd uit opgeslagen weefsel uit het pathologie archief, voor een specifieke genetische analysetechniek, namelijk de Drug Metabolizing Enzymes and Transporters (DMET) Plus array. Deze array is gebruikt voor verschillende farmacogenetische studies in dit proefschrift. Aangezien een deel van de geïncludeerde osteosarcoma-patiënten niet meer in leven was op het moment van inclusie, was van deze patiënten geen bloed of speeksel beschikbaar voor DNA isolatie, maar opgeslagen weefsel (formaline gefixeerd en ingebed in paraffine) bestaande uit tumor met omringend normaal weefsel. In het algemeen is het DNA verkregen uit dergelijk materiaal van mindere kwaliteit, wat invloed kan hebben op de resultaten bij genetische analyses. We hebben daarom de genetische resultaten vergeleken tussen kiembaan DNA geseleerd uit normaal opgeslagen weefsel en bijbehorende bloed- of speekselsel monsters van zestien osteosarcoma-patiënten. We hebben daarbij een hoge overeenstemming (97.4%) gevonden tussen de resultaten verkregen uit opgeslagen weefsel en bloed of speeksel. Dit resultaat geeft aan dat DNA
De hoofdstukken 5-7 beslaan onderzoek dat gericht is op de farmacogenetica van de effectiviteit van chemotherapie bij osteosarcoompatiënten. In hoofdstuk 5 beschrijven we een studie waarin we de reproduceerbaarheid hebben onderzocht van een door onze groep eerder gevonden associatie van een genetische risico score, gebaseerd op enkele farmacogenetische markers, met 5-jaars progressievrije overleving (PFS). De genetische risicoscore is het aantal ongunstige genetische varianten (risico allelen) van een patiënt. Hierbij maakten we gebruik van zowel een Nederlands, Spaans als Engelse cohort. Hoewel een van de varianten (in CYP3A4) niet betrouwbaar bepaald kon worden, konden we de associatie van de genetische risicoscore op basis van markers in FastL, MSH2, CASP3 en ABCG5 bevestigen. Patiënten zonder of met één risico allel hadden een 5-jaars PFS van 78.5%, vergeleken met een 5-jaars PFS van 59.5% in patiënten met 2-3 risico allelen, 48.0% met 4 of 5 risico allelen en 57.1% met meer dan 5 risico allelen (P<0.040). Daarnaast konden we de associatie van twee individuele genetische varianten (in MSH2 en ABCG5) met 5-jaars PFS bevestigen in een meta-analyse waarin ook de cohorten van de oorspronkelijke studie werden meegenomen. Deze farmacogenetische bevindingen zijn mogelijk van nut voor stratificatie van patiënten bij diagnose en kunnen een basis vormen voor farmacogenetische screening in de toekomst.

Hoofdstuk 6 beschrijft een volgende stap in het onderzoek naar farmacogenetische markers voor de effectiviteit van chemotherapie, om (additionele) associaties te vinden van varianten in 231 ADME genen met 5-jaars PFS middels een DMET analyse. Na analyse in een discovery fase met twee Nederlandse osteosarcoomcohorts (totaal 283 patiënten) en validatie in Spaanse en Engelse cohorten, identificeerden we acht genetische varianten in relatie tot 5-jaars PFS. Van deze varianten was rs17064 in het ABCB1 gen onafhankelijk gevalideerd. Zeven andere varianten in vijf genen coderend voor transporters (ABCC4, ABCC5, ABCG1, SLC22A5, SLC22A14) waren significant geassocieerd met 5-jaars PFS in een gecombineerde meta-analyse van analyse- en validatiecohorts. Deze bevindingen dragen bij aan de aanwijzingen uit andere studies dat varianten in ABCB1 van invloed zijn op de uitkomst na behandeling van osteosarcomen met chemotherapie. Daarnaast geeft het voor het eerst aan dat ook varianten in genen coderend voor solute-linked carriers betrokken zijn bij de respons op chemotherapie in osteosarcoompatiënten. Deze markers zijn mogelijk van belang bij het ontwikkelen van nieuwe behandelstrategieën en optimalisatie van de huidige behandeling van patiënten met een osteosarcoom.

Hoofdstuk 7 beschrijft tevens een DMET analyse, waarbij we hebben gefocust op de subgroep van patiënten met een inadequate respons op de (chemotherapeutische) behandeling, gedefinieerd als progressieve ziekte tijdens of kort na afloop van de behandeling. Bij analyse in een Nederlands discoverycohort van 283
osteosarcoom-patiënten, gevolgd door validatie in een gecombineerd Spaans en Engels cohort en tevens Australisch cohort, vonden we een onafhankelijk gevalideerde associatie van een variant in het SLC7A8 gen (rs1884545) met progressieve ziekte, waarbij de genetische variant een beschermend effect liet zien. Daarnaast waren drie varianten in andere genen (CYP8B1, SLC22A2, SLC22A5) significant geassocieerd met progressieve ziekte in een meta-analyse van de resultaten van discovery- en validatiecohorts. Deze exploratieve analyse is een eerste stap om patiënten te kunnen identificeren voor wie de behandeling met chemotherapie niet effectief is.

In hoofdstuk 8 zijn de belangrijkste bevindingen van het onderzoek zoals in dit proefschrift beschreven samengevat, met daarbij een discussie van de sterke punten en beperkingen van het onderzoek. Daarnaast wordt een mogelijk toekomstig perspectief gegeven op de behandeling van osteosarcomen en het (farmacogenetisch) onderzoek in en voor deze patiënten.
Hanneke Vos was born on the 13th of August 1986 in Delft, The Netherlands. In 2004, she obtained her VWO diploma at the Christelijk Lyceum Delft in Delft. She started with the Bachelor Biomedical Sciences at Leiden University in 2005 after a gap year. During her Bachelor studies, she did a research internship at the Department of Hemostasis and Thrombosis (Eindhoven laboratory) at the Leiden University Medical Center (LUMC), where she focused on the role of tissue factor signal transduction in tumor progression (Prof. dr. H.H. Versteeg). After completing the Bachelor, she did her Master studies at the same university. Her first Master research internship took place at the Laboratory of Immunology, Department of Pediatrics of the LUMC, where she studied histone deacetylase treatment for Ewing sarcoma (Dr. D. Berghuis, Dr. M.W. Schilham, Dr. A.C. Lankester). She finished her Master studies with a research internship and thesis on the effect of chemotherapeutic treatment on the immune system in advanced cervical cancer patients at the Centre for Human Drug Research in Leiden (Dr. H. van Meir and Prof. dr. J. Burggraaf), where she was appointed as (replacing) project leader during the last three months. After obtaining her Master’s degree in 2011, she started her PhD project on the pharmacogenetics of treatment response in osteosarcoma at the Laboratory of Pediatric Oncology (Dr. F.N. van Leeuwen), Department of Pediatrics of the Radboud University Medical Center, Nijmegen, The Netherlands in 2012. The project was supervised by Dr. D.M.W.M. te Loo (Pediatric Oncology), Dr. M.J.H. Coenen (Human Genetics) and Prof. dr. H.J. Guchelaar (Pharmacology and Toxicology, LUMC). The results of the research carried out during this PhD project are presented in this thesis.
List of publications


* equal contribution
Dankwoord

Wat ben ik blij dat ik in dit proefschrift vier jaar onderzoekswerk heb kunnen bundelen tot een mooi eindresultaat. Een groot aantal mensen heeft een bijdrage geleverd aan mijn promotietijd en op deze plaats wil ik hen van harte bedanken.

Dank aan…

…alle patiënten (en hun ouders) die bereid waren om deel te nemen aan het onderzoek.

…mijn promotoren. Kees, dank voor je rol bij mijn promotietraject en voor je vertrouwen. Henk-Jan, je positieve betrokkenheid, kritische blik en snelle feedback op artikelen waren zeer waardevol.

…mijn copromotoren voor jullie begeleiding en vertrouwen. Maroëska, dank voor het bieden van deze promotieplek, natuurlijk je onmisbare klinische blik en bovenal je niet afhaalbare optimisme en aanstekelijke enthousiasme. Dank ook voor de kansen die je gaf om naar grote congressen te gaan, waarbij de trips naar Chicago, maar ook de workshops en congresdagen, waarbij de gelegenheid om te sparen. Dank ook voor de gezellige (en effectieve) gezamenlijke schrijfuren, wat gaf dat een energie. Jouw kennis en ervaring in de (farmaco)genetica waren onmisbaar en jouw nuchtere aanpak en besluitvaardigheid waren voor mij een voorbeeld.

…de leden van de manuscriptcommissie, Prof. Van der Hoeven, Prof. Kiemeney en Dr. Merks, voor het beoordelen van dit proefschrift.

…Frank, voor het bieden van de plek in de LKO groep en voor je betrokkenheid, interesse en adviezen bij de voortgang en inhoud van het onderzoek, ook al werd het geen celbiologisch werk.

…alle collega’s. Ik wil in het bijzonder de LKO groep (Esther, Kirsten V, Jern, Yuni, Miriam, René, Liesbeth, Dorette, Marc, Laurens en Blanca) bedanken voor jullie betrokkenheid, goede tips en gezelligheid en daarvan Jeroen, Arthur en Agata voor de leuke tijd op kantoor in het Trigon! Daarnaast wil ik ook de collega’s van het LKI bedanken, in het bijzonder de kamergenoten sinds de verhuizing naar het ziekenhuis (Erika, Ria, Kirsten, Marrit, Inge, Corné, Jop, Fredrick, Esther W), het was fijn werken met jullie gezelligheid, hulp en goede gesprekken! Ada, geweldig hoe je altijd klaar stond en alles wist te regelen. Dank ook aan de collega’s van het MF lab. Hoewel ik niet bij jullie mijn vaste plek had, was er altijd een warme sfeer, gezelligheid op het lab en waardevolle input bij werkbesprekingen.

Dank voor, en voor het vele genotyperingswerk dat jullie hebben verzet.

…Melanie, voor je betrokkenheid en het vele werk waar ik op voort kon bouwen.

…alle coauteurs, voor de samenwerking, alle inspanningen, adviezen en input voor het onderzoek en de artikelen.

…iedereen die me in de verschillende UMC’s heeft geholpen (patiënt)gegevens en weefsels boven water te halen, van Amsterdamse datamanagers tot Nijmeegse (pathologie) archieffinanciers en Leidse (poli)secretarissen.

…de studenten die aan dit proefschrift hebben bijgedragen door literatuuronderzoek en een onderzoeksproject. Chanel en Marije, ik heb jullie met plezier begeleid.


…mijn paranormen, superleuk dat jullie bij de verdediging naast me staan! Erika, ik ben heel blij met de leuke vier jaren van samen werken, eten, sporten, lunchwandelrondes, theepauzes en elkaar helpen in de uitdagingen van het PhD bestaan. Marloes, zoals al heel lang heb ik ook de afgelopen jaren genoten van onze vriendschap, de leuke uitrusting en het bijkletsen over werk, promoveren en al het andere.

…mijn (schoon)familie. Lieve pap en mam, super bedankt dat jullie er altijd zijn met zo veel liefde en betrokkenheid. Jullie steun, wijze woorden en gezelligheid betekenen veel voor me! Lieve Marieke en Tim, Gert-Jan en Sherise, het is altijd heel fijn om jullie te zien en te spreken, dichtbij of ver weg! Anneke, dankjewel voor je gastvrijheid en gezelligheid de eerste drie maanden in Velp. En wat ben ik blij met ook de lieve (bijna officiële) schoonfamilie, Paul en Carla, Norbert en Iris en de meiden, dank voor alle gezelligheid en betrokkenheid.

…Patrick, mijn lief voor het leven! Jou leren kennen was het absolute hoogtepunt van de afgelopen vier jaar en ik ben ontzettend dankbaar voor zoveel liefde en support. Je bent geweldig en ik geniet van het leven samen. Een supermooi nieuw hoofdstuk samen staat voor de deur!