Aus der Klinik für Allgemein-, Viszeral-, und Transplantationschirurgie

Klinikum der Ludwig-Maximilians-Universität München

Establishment and validation of comprehensive prognostic models in gastric cancer and sarcoma: the integration of clinicopathological factors and gene signatures

Dissertation

zum Erwerb des Doktorgrades der Medizin

an der Medizinischen Fakultät der

Ludwig-Maximilians-Universität München

vorgelegt von

Haoyu Ren

aus

Chongqing, China

Jahr

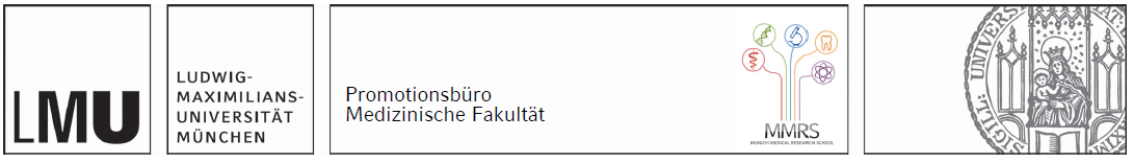
2022

Mit Genehmigung der Medizinischen Fakultät der

Ludwig-Maximilians-Universität zu München

Erster Gutachter:	Prof. Dr. med. Martin Angele
Zweiter Gutachter:	Prof. Dr. Dr. Jens Neumann
Dritter Gutachter:	Prof. Dr. Axel Imhof
Mitbetreuung durch den promovierten Mitarbeiter:	PD. Dr. med. Florian Bö sch
Dekan:	Prof. Dr. med. Thomas Gudermann
Tag der mündlichen Prüfung:	02.06.2022

Affidavit



Affidavit

Ren Haoyu
_____
Surname, first name

Guardinistraße 139
_____
Street

81375, Munich, Germany
_____
Zip code, town, country

I hereby declare, that the submitted thesis entitled:

Establishment and validation of comprehensive prognostic models in gastric cancer and sarcoma: the integration of clinicopathological factors and gene signatures

……

is my own work. I have only used the sources indicated and have not made unauthorised use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the dissertation presented here has not been submitted in the same or similar form to any other institution for the purpose of obtaining an academic degree.

Munich, 07.01.2021
_____
place, date

Haoyu Ren
_____
Signature doctoral candidate

Table of content

Affidavit **3**

Table of content **4**

List of abbreviations **5**

List of publications **7**

Contribution to the publications **8**

1.1 Contribution to paper I 8

1.2 Contribution to paper II 8

1.3 Contribution to paper III 8

2. Introduction **9**

2.1 Prognostic factors in cancer 9

2.2 Prognostic factors and biomarkers in gastric cancer 10

2.3 Prognostic factors and biomarkers in sarcoma 12

2.4 The rise of transcriptome in the age of precision oncology 14

2.5 The applications of gene expression signatures in tumor prognosis 15

2.6 The application of nomograms and decision trees in cancer management 17

2.7 The predictive function of gene signatures for therapeutic response 18

2.8 Aims of the present studies 19

3. Summary (in Englisch) **21**

3.1 Paper 1: 21

3.2 Paper 2: 21

3.3 Paper 3: 22

4. Zusammenfassung (deutsch) **23**

4.1 Paper 1: 23

4.2 Paper 2: 23

4.3 Paper 3: 24

5. Paper I **26**

6. Paper II **27**

7. Paper III **28**

References **29**

Acknowledgements **35**

List of abbreviations

ARGs	Angiogenesis-related genes
ARMS	Alveolar rhabdomyosarcoma
AUC	Area under curve
BC	Breast cancer
CART	Classification and regression trees
CHFR	Checkpoint with forkhead and ring finger domains
CRC	Colorectal cancer
DEGs	Differentially expressed genes
DFS	Disease-free survival
DT	Decision tree
ECM	Extracellular matrix
EGFR	Epidermal growth factor receptor
EMT	Epithelial-mesenchymal transition
GC	Gastric cancer
GSEA	Gene set enrichment analysis
HER2	Human epidermal growth factor receptor 2
HGF	Hepatocyte growth factor
ICGC	International Cancer Genome Consortium
ImmuCellAI	Immune Cell Abundance Identifier
IRGs	Immune-related genes
LNs	Lymph nodes
MLA	Machine learning algorithm
MMPs	Matrix metalloproteinases
MPNST	Malignant peripheral nerve sheath tumors
MSI	Microsatellite instability
OS	Overall survival
PCNA	Proliferating cell nuclear antigen
RNA-Seq	RNA-sequencing
SEER	Surveillance, Epidemiology, and End Results
STS	Soft tissue sarcoma
TCGA	The Cancer Genome Atlas
TFG-β1	Transforming growth factor-β1

TME	Tumor microenvironment
TNM	Tumor, Node, Metastasis
VEGF	Vascular endothelial growth factor
WGCNA	Weighted gene co-expression network analysis

List of publications

1. **Ren H**, Zhu J, Yu H, Bazhin AV, Westphalen CB, Renz BW, Jacob SN, Lampert C, Werner J, Angele MK, Bösch F. Angiogenesis-Related Gene Expression Signatures Predicting Prognosis in Gastric Cancer Patients. Cancers (Basel). 2020 Dec 8;12(12):3685. doi: 10.3390/cancers12123685. PMID: 33302481

2. **Ren H**, Bazhin AV, Pretzsch E, Jacob S, Yu H, Zhu J, Albertsmeier M, Lindner LH, Knösel T, Werner J, Angele MK, Bösch F, A novel immune-related gene signature predicing survival in sarcoma patients, Molecular Therapy: Oncolytics (2022), doi: https://doi.org/10.1016/j.omto.2021.12.007.

3. Jacob S, Jurinovic V, Lampert C, Pretzsch E, Kumbrink J, Neumann J, **Haoyu R***, Renz BW, Kirchner T, Guba MO, Werner J, Angele MK, Bösch F. The association of immunosurveillance and distant metastases in colorectal cancer. J Cancer Res Clin Oncol. 2021 Nov;147(11):3333-3341. doi: 10.1007/s00432-021-03753-w. Epub 2021 Sep 2. PMID: 34476575

*：*Unfortunately, my name was accidentally reversed when we published the paper, it should be "Ren H"*.

Contribution to the publications

1.1 Contribution to paper I

Angiogenesis-Related Gene Expression Signatures Predicting Prognosis in Gastric Cancer Patients.

First author: conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing—original draft preparation, and writing—review and editing.

1.2 Contribution to paper II

A novel immune-related gene signature predicting survival in sarcoma patients.

First author: project design, technical and methodological support, statistical analysis, writing—original draft preparation, and writing—review and editing.

1.3 Contribution to paper III

The association of immunosurveillance and distant metastases in colorectal cancer.

Co-author: formal analysis, methodology, validation.

2. Introduction

Cancer refers to a disease characterized by the uncontrolled growth and division of abnormal cells that have the ability to invade other organs. In 2020, about 19.3 million new cancer cases are expected to be diagnosed worldwide with about 10 million patients expected to die of cancer [1]. During the last decades, various clinical and pathological characteristics have been introduced and evaluated to assess the survival outcome of patients with malignant tumors.

In the case of gastric cancer (GC) different evaluation systems have been established improving therapy of GC, such as the AJCC TNM stage, Lauren classification and Borrmann classification. However, these traditional analyses of cancer were mainly based on histological morphology and location of growth. These routine methods do not take into account the individual biologic features of the patient's tumor. Previous studies have revealed that the conventional Tumor, Node, Metastasis (TNM) staging systems have limited value in specific cancer patients [2-4]. Furthermore, even tumors in the same stage could be still highly heterogeneous, which would exhibit different survival rates and treatment responses [5, 6].

In clinical practice, TNM staging still occupies a core position in the current prognosis evaluation system. As the new and promising prognostic factors genetic biomarkers and gene signatures have been reported, but TNM provides a solid foundation to establish risk stratification. Hence, innovative approaches should be developed to help increase overall prognosis without losing the vital anatomic content of TNM. In recent years, the application of artificial intelligence and machine learning algorithms (MLA) in medicine provided an opportunity to integrate numerous and various prognostic factors into a system that could surpass traditional staging improving the prediction of outcome [7].

2.1 Prognostic factors in cancer

To better manage individual patients with cancer, the prognostic factors in cancer should be well recognized. In 2003, Gospodarowicz et al. proposed three broad groupings of prognostic factors: tumor-related prognostic factors, host-related prognostic factors, and environment-related prognostic factors [8].

The most common tumor-related prognostic factors include TNM stage, tumor histopathology, and tumor biology, which are directly associated with the tumor itself. TNM stage describes anatomic disease extent, including depth of invasion and size of the tumor, the status of lymph node metastasis, and the presence of distant metastasis. In addition, numerous molecular biomarkers reflecting tumor biological behaviors have also been shown to influence patient's prognosis for a variety of cancers.

Host-related prognostic factors refer to any basic information of patients, such as age [9], gender [10], and ethnicity [11]. Furthermore, other important parameters, such as the physical state, immune function [12], nutritional issues [13], and comorbid conditions [14], also have a considerable impact on tumor prognosis. Although most of them are not directly related to tumors, these parameters might affect patient's response to the tumor.

Environment-related prognostic factors mean the external factors acting on the patients, including types of treatment methods, insurance status, and income level [15]. These factors have been shown to influence the outcome of patient's disease, especially in developing countries.

The goal of the present studies is to establish innovative and comprehensive prognostic prediction models in two types of malignancies of different origins (GC and sarcoma). As aforementioned, clinicopathological factors remain the cornerstone of prognosis evaluation. In addition, new and promising molecular biomarkers will be developed to enhance the prediction capability.

2.2 Prognostic factors and biomarkers in gastric cancer

GC is a major cause of death worldwide, with 5-year survival rates reaching only approximately 35-45% [16, 17]. The identification of prognostic factors is a significant aspect of cancer management of risk stratification and treatment strategies. The prognostic factors and biomarkers for GC currently being used will be discussed in more detail below.

2.2.1 Clinicopathological prognostic factors of gastric cancer

In 1998, the ten-year results of the German gastric cancer study showed that the nodal status, the pathological T stage, and the presence of postoperative complications were the major independent prognostic factors [18]. Among them, depth of invasion and lymph node metastasis have been proved in further studies to be the most significant prognostic variables [19, 20]. Patients with an invasion of the mucosa and submucosa have a 5-year survival rate of 89.8-93.4%. However, when the tumor invades the subserosa or the serosa, the survival rates dropped to 60.5% or 39.7%, respectively [20]. The status of lymph node metastasis is also regarded as an important predictor for patient survival. The patients with the number of lymph nodes (LNs) involved is $\geqslant 4$ or the ratio of involved-to-resected LNs exceeds 0.5 have a significantly shorter survival time [19, 20]. In addition, gross type, location of primary tumor and histological differentiation have been confirmed to affect the prognosis of patients moderately [19, 21, 22].

However, the TNM staging based on the prognostic clinicopathological parameters still has some limitations for specific patients. For example, Zhao et al. proposed to adjust the 8^{th} edition of TNM staging, in which T4aN2 was added into stage IIIB, T4aN3a was incorporated into stage IIIC, and T4bN3b was classified as stage IV. This reclassification can provide better risk stratification than original staging [23]. Additionally, another problem representing a bias is that data collection for the staging of GC has focused on the patients after surgery, while patients treated with preoperative neoadjuvant chemotherapy were not included [24]. Due to frequent changes of tumor biological characteristics during treatment, it is not enough to obtain an accurate outcome prediction by TNM staging alone. Hence, it is necessary to further improve the current prognosis prediction system that can better reflect tumor heterogeneity.

2.2.2 Classical molecular biomarkers in gastric cancer

Since invasion and metastasis are critical factors affecting the mortality of GC, numerous genes and molecules involved in these processes have been confirmed to be powerful prognostic indicators [25]. Here, the most common prognostic molecular biomarkers of GC were summarized briefly.

GC cells produce a variety of growth factors and corresponding receptors to generate biological functions, including cell growth, angiogenesis, and extracellular matrix degradation for tumor invasion and metastasis. Multiple studies have confirmed that high expression levels of epidermal growth factor receptor (EGFR) and human epidermal growth factor receptor 2 (HER-2) predicted worse overall survival (OS) of GC patients [26]. These receptors induce various downstream pathways, promoting cell proliferation, migration, angiogenesis, and apoptosis inhibition [27]. In respect to angiogenesis, vascular endothelial growth factor (VEGF) is thought to be the most important factor driving tumor angiogenesis. Previous studies showed that high VEGF levels were associated with lymph node metastasis or distant metastasis and thereby predicted a worse prognosis [28]. Moreover, the expression of transforming growth factor-β1 (TFG-β1) can promote GC progression by indirectly triggering neovascularization through the up-regulation of VEGF levels [29]. Tumor invasion through extracellular matrix (ECM) is modulated by matrix metalloproteinases (MMPs) which induce ECM degradation and create paths for migration. In GC, MMP-2, MMP-7, and MMP-9 have been viewed as significant biomarkers predicting poor survival [30-32]. In addition, other molecular and genetic markers have also been confirmed to have prognostic significance, such as cell-cycle regulators (cyclin G2 [33], cyclin D2 [34], p27 [35], and p53 [36]), telomeres and telomerase (hTERT [37] and POT1 [38]), and cell-adhesion markers (E-cadherin [39], CD44 [40], and CD44v6 [41]).

Defects in the DNA mismatch repair were observed in 10-22% of GC, leading to genomic instability characterized by microsatellite instability (MSI) [42]. The current evidence showed that patients with MSI-high GC had a better prognosis than MSI-low GC patients because MSI-high tumors were associated with less prevalent lymph node metastasis [43-45]. Moreover, the results of the MAGIC trial indicated that patients with operable MSI-high GC might not benefit from perioperative chemotherapy [46]. In this respect, MSI is a promising prognostic indicator for survival outcome and treatment decision.

In GC, epigenetic alterations such as DNA methylation and histone modifications can regulate gene expression and thereby affect tumor progression. DNA methylation changes the readability of the DNA and leads to the inactivation of a gene by inhibiting transcription [47]. The most common methylation regions are CpG islands, which are correlated with the silencing of multiple tumor suppressor genes and tumorigenesis process. These genes are inter alia CDH1, hMLH1, p16INK4a, RAR-beta, MGMT. The study from Oue et al. found that accumulation of DNA methylation may contribute to progression of most GCs. Recently, several DNA methylation signatures were developed to predict the prognosis of gastric cancer patients and the performance of predictive models based on DNA methylation are favorable [48-50].

2.3 Prognostic factors and biomarkers in sarcoma

Sarcomas are a rare and heterogeneous group of tumors originating from stromal cells, including more than 60 subtypes with different biological and clinical features [51]. They can be divided into two general categories, soft tissue sarcoma (STS), and primary bone sarcoma. In general, the 5-year survival rate of sarcoma patients is only about 50% due to high incidence of distant metastasis [52, 53]. Previous studies have identified useful prognostic variables of sarcomas for local recurrence, distant metastasis, and tumor-related mortality, but some predictive factors remain controversial with many opposing reports [54]. Reassuringly, molecular profiling of sarcomas has revealed the deeper biological mechanism of these rare tumors and pinpointed novel molecular biomarkers, which opens novel avenues for precision medicine in the field of sarcoma.

2.3.1 Clinicopathological prognostic factors of sarcoma

A study of 389 patients with STS showed that tumor location, resection margins, size, age at diagnosis, and tumor grade were significant prognostic factors for tumor relapse or survival outcome [55]. Compared with STS in the head, neck and trunk, the extremity lesions had a better local control and a longer survival time [55]. Currently, there is a general consensus that the histologic grade of STS is the most important prognostic factor for local recurrence and OS [56, 57]. The high-grade (G3/G4) STS patients exhibited a worse survival than patients with low-grade (G1/G2) tumors because poor differentiated tumor cells are more aggressive [57]. The certain correlation between tumor size and prognosis has been investigated well. For trunk, extremities, and retroperitoneal STS, 5cm and 10cm are two critical prognostic cut-off values according to the current staging systems [58]. The resection margins in STS surgery are also viewed as a risk factor for local recurrence and metastasis. A previously published study showed that the 5-year local control rate of patients with negative margins reached 78% while it decreased to 52% in patients with positive margins [59].

In addition, different bone sarcomas also have homogenous prognostic factors. A Surveillance, Epidemiology, and End Results (SEER) data-based study showed that older age (>40), higher grade, and advanced stage were significantly correlated with worse prognosis in patients with bone sarcoma [60]. Furthermore, different tumor sites also affected patient's survival. Some studies reported osteosarcomas in the spine and pelvic were linked with an unfavorable prognostic outlook [61, 62]. Although the identification of these prognostic factors are validated tools for tumor staging, the current evaluation system merits further improvement due to high tumor heterogeneity.

2.3.2 Biological prognostic factors of sarcoma

In the last decade, numerous novel molecules have been identified in different sarcoma subtypes, which have added to the wealth of information that may be predictive of biological behavior and

prognosis of sarcomas. Here, representative molecular biomarkers developed for sarcoma in recent years were described shortly below.

Proliferative activity plays an important role in determining the prognosis of sarcoma patients. The most common two parameters are the cell proliferation markers Ki-67 and proliferating cell nuclear antigen (PCNA). High expression rates of Ki-67 and PCNA have been associated with a reduced OS in a series of STS [63, 64], synovial sarcoma [65, 66], and primary osteosarcoma [67, 68]. Additionally, DNA ploidy analysis and S-phase rate could also indicate the status of cell proliferation. Several studies found robust prognostic power for the measurement of DNA ploidy patterns and S phase fraction through DNA flow cytometry [69-71].

The role of oncogenes in sarcomas has been extensively investigated. The tumor suppressor p53 play an important role in mediating cell response to diverse stresses and the function of p53 is often altered in cancers. P53 overexpression in the nucleus is regarded as a surrogate marker for p53 mutation. Multiple studies have confirmed that p53 was an independent prognostic biomarker for STS [72], gastrointestinal sarcoma [73], Ewing's sarcoma [74], synovial sarcoma [75], and osteosarcoma [76]. Moreover, MDM2, encoded by the MDM2 gene, is the main regulator of p53 and has a synergistic effect with p53 in the progression of sarcoma [77, 78].

C-Myc is frequently deregulated in human cancers. The expression of c-Myc has been correlated with poor prognosis in several types of sarcomas, including synovial sarcoma [79], osteosarcoma [80], Ewing's sarcoma [81], liposarcoma [82], and soft tissue leiomyosarcoma [83].

The HER-2/neu oncogene (also known as c-erbB-2), encodes a transmembrane protein (p185) which is structurally homologous to EGFR. Several reports have analyzed the correlation between the HER-2 expression and the prognosis of patients with synovial sarcoma. One study showed that patients with high HER-2/neu mRNA levels had a longer disease-free survival (DFS) than those with low HER-2/neu mRNA levels, while the other study found no definite correlation between HER-2/neu expression and survival outcome [84, 85]. Since the case numbers of these two studies are small, further clinical investigations are needed.

In specific types of sarcomas, some gene fusions generated by chromosomal translocation have been shown to be prognostically relevant. For example, EWS-FLI1, a tumor-specific chimeric transcription factor, is the most predominant fusion in Ewing's sarcoma (85% of cases). Two studies demonstrated that type 1 EWS-FLI1 was a significant predictor of favorable prognosis of patients with Ewing's sarcoma [86, 87]. Cytogenetic studies have found a characteristic SYT-SSX fusion gene in almost all synovial sarcomas, usually SYT-SSX1 or SYT-SSX2. Survival analyses from previous two studies has revealed that patients with SYT-SSX2 had a significantly longer metastasis-free survival than patients with SYT-SSX1 [88, 89].The reason for this difference may be that SYT-SSX1 type fusion was highly correlated with high tumor cell proliferative activity [89]. In addition, PAX3-FKHR or PAX7-FKHR gene fusions were identified in over 70% of alveolar rhabdomyosarcoma (ARMS) patients. In ARMS patients with metastasis, PAX3-FKHR fusion is a marker of poor prognosis [90].

Previous studies have also investigated the prognostic value of cell adhesion molecules in sarcomas. Saito et al. have shown that the reduced expression level of E-cadherin and α-catenin and aberrant nuclear β-catenin expression predicted a poor OS in synovial sarcoma patients [91]. Dysadherin, a cancer-associated membrane glycoprotein, inhibits E-cadherin expression and promotes tumor metastasis. The high dysadherin expression level was an independent adverse prognostic factor in epithelioid sarcoma [92] and synovial sarcoma [93]. Like other solid tumors, several growth factors and their associated receptors play a critical role in the occurrence and

progression of sarcoma. The overexpression of EGFR is confirmed to be a significant poor prognostic marker of adult STS, which is highly correlated with the histologic grade [94]. This result suggests that some STS patients could benefit from the treatment with inhibitors of EGFR.

The hepatocyte growth factor (HGF)/c-MET autocrine signaling has been implicated in the carcinogenesis of various tumors. A study of 69 cases of synovial sarcoma showed that the co-expression of HGF and c-MET was associated with higher proliferative activities and worse survival outcome [95]. Numerous observations have indicated that abnormalities of mitotic checkpoint genes are closely related to chromosomal instability and tumorigenesis [96, 97]. The checkpoint with forkhead and ring finger domains (CHFR) gene, a mitotic checkpoint gene, has been identified as a tumor suppressor in malignant peripheral nerve sheath tumors (MPNST). The reduced expression of CHFR was a significantly poor prognostic factor in MPNST [98].

In addition to the above mentioned prognostic factors, there are other biomarkers reported to have prognostic impact in sarcoma patients, including the cell cycle regulator p21 [99], multidrug resistance related genes ABCB1 [100] and YB-1 [101], and the tumor suppressor RASSF1A [102].

2.4 The rise of transcriptome in the age of precision oncology

Recent achievements in transcriptomics provided an unprecedented opportunity to improve cancer management. Transcriptomics is the analysis of the complete set of RNA transcripts that are generated by the genome using high-throughput techniques. The transcriptomic data can be viewed as a snapshot of the temporary cell state. It not only profiles genomic backgrounds to establish a global picture of cell function, but also analyze actively expressed genes and transcripts under different conditions [103].

Various technologies have been used to analyze the transcriptome, including hybridization-or sequence-based methods. Currently, gene expression microarray and RNA-sequencing (RNA-Seq) stand out for transcriptome analysis almost as a routine method, which has been widely used in biological, drug and clinical research.

Gene expression microarray technology (also known as GeneChip) refers to the fixation of a large number of probes molecules on the solid surface and hybridization with labeled RNA or cDNA. This technology enables measuring the expression levels of millions of genes simultaneously. With the recent advent of microarrays, searches for novel genes involved in carcinogenesis and better biomarkers for prognosis have been well established. For example, D'Errico et al. analyzed the transcriptional profile of 38 GC patients with respect to MSI using oligonucleotide chips. The authors found that distinct immune and apoptotic related genes can efficiently discriminate between MSI-high and MSI-low tumors [104]. In addition, molecular subtyping based on microarrays has been widely used in various cancers, including breast [105], lung [106], liver [107], colon [108], stomach [109] and soft tissue sarcomas [110]. Furthermore, a previous study regarding malignant gliomas demonstrated that a gene expression-based classification correlated more precise with prognosis than a histomorphologic pathological classification.

RNA-Seq is a sequencing-based approach for the detection and quantitative analysis of RNA in a sample. Compared with microarrays, RNA-seq shows numerous advantages for detecting novel transcripts, single-nucleotide variants, and splice junctions [111]. Moreover, RNA-Seq has a low

background signal and consequently quantify gene expression across a larger dynamic range. Due to the high costs and high data-storage requirements of RNA-Seq, large-scale RNA-Seq transcriptome analyses of various types of cancers were collected only by some large databases, such as The Cancer Genome Atlas (TCGA) and International Cancer Genome Consortium (ICGC).

In summary, next generation sequencing generates numerous valuable genomic data of cancer samples. Meanwhile, various advanced MLA have also been designed for processing and translation of such an amount of data [112]. There is evidence that the combination of high-throughput transcriptomic profiling and MLA will efficiently and accurately improve clinical practice.

2.5 The applications of gene expression signatures in tumor prognosis

In the last decades, hundreds of molecular prognostic and predictive markers in various types of cancers have been identified. However, these proposed biomarkers have not the expected impact on clinical decision making. Because of unavoidable tumor heterogeneity, it is difficult to predict survival outcome accurately according to a single biomarker or factor. Thus, identifying subsets of prognostic genes and establishing gene expression signatures or prognostic models may lead to new and more precise clinical strategies.

A gene expression signature includes a list of genes that jointly present robust predictive performance for a survival event or classification. Generally, signature genes are decided based on statistical methods, including Least absolute shrinkage and selection operator (LASSO) regression analysis, Cox's proportional hazard and regression model, or MLA [113]. Currently, several gene expression signatures have been established to optimize survival prediction and to identify which part of cancer patients need additional treatment. For example, the 21-gene recurrence score is the most common commercial multigene assay in breast cancer (BC), which not only predict the prognosis of early-stage BC but also identify high-risk patients who might benefit from additional chemotherapy [114]. This panel contains 16 cancer-related genes and 5 reference genes. The values of gene expression are converted into a recurrence score through a proprietary algorithm, which ranges from 0 to 100 and a higher score indicates a greater probability of disease recurrence. Three different risk level subgroups (low-, intermediate-, and high-risk) were identified and demonstrated a clear disparity in survival outcomes: the relapse-free interval and OS ($P<0.001$ for both) [115]. In the last decade, several clinical trials have been conducted to validate the prognostic utility of the 21-gene recurrence score. The remarkable findings from these studies further at least in part paved the inclusion of the 21-gene signature in major treatment guidelines for breast cancer [116].

The great success of multigene assay for breast cancer has set a benchmark in the field of clinical cancer research. Although there is no such gene signature incorporated into clinical routine practice in other malignant tumors, a large number of gene expression profiling investigations on tumor prognosis have been reported. In general, there are three major strategies to establish gene signatures [117]. The first strategy is screening for differentially expressed genes (DEGs) and identifying the final candidate genes through statistical approaches, such as Cox regression and LASSO regression analysis. The second strategy focus on one gene set of specific pathways or

biological functions, like metastasis-related genes and immune-related genes, then perform dimensionality reduction and modeling. The third one refers to construct the gene signatures based on a specific gene family, such as NDRG family and m6A regulatory gene family.

In GC, numerous prognostic gene signatures have been developed, most of which are based on TCGA and Gene Expression Omnibus (GEO) datasets. For example, Cho et al. established a 6-gene (CTNNB1, EXOCS3, TOP2A, LBA1, CCL5, and LZTR1) risk scoring system by which patients with a higher risk score have a significantly lower DFS rate in three different GC cohorts [118]. It is increasingly definite that immune and stromal cells in the microenvironment are critical to cancer progression [119]. Hence, Wang et al. found stromal-immune score-related DEGs and constructed 4-gene signature (SOX9, LRRC32, CECR1, and MS4A4A) as a prognosis stratification tool for GC patients in both TCGA and GEO cohorts [120]. Furthermore, a study focusing on epithelial-mesenchymal transition (EMT) proposed a novel EMT-related gene signature (ITGAV, DAB2, SERPINE1, MATN3, PLOD2) with the 5-year OS area under curve (AUC) value of 0.784, indicating a good prognostic discrimination in GC patients [121].

In contrast to epithelial tumors, sarcomas occur in all organs of the body, so histological grade and clinicopathological characteristics form the core prognostic criteria. In 2010, Chibon et al. proposed a complexity index in sarcomas including 67 genes related to mitosis and chromosome instability. Significantly, it not only predicted metastasis events in the STS cohort, but also showed powerful predictive ability in other cancers, such as gastrointestinal stromal tumors, BC, and lymphomas [122]. Since sarcomas contain multiple histological subtypes, most gene expression signatures are developed based on a specific subtype. For example, Ren et al. constructed an 11-gene signature based on immune-related genes for Ewing sarcoma, which present a high prognostic value in two independent cohorts [123]. In addition, Shi et al. focus on metastasis-related genes and established a 3-gene (MYC, CPE, and LY86) signature for osteosarcoma patients, demonstrating the robust prediction efficiency of OS [124].

Despite these great strides, it has also become clear that only a few gene signatures have reached the stage of clinical utility so far. It is largely because most studies on prognostic gene expression signature are retrospective in nature. Bias and confounding are more common in these studies, leading to a lack of reproducibility of the results. Thus, further clinical investigations based on high-level evidence are needed. Secondly, the vast majority of gene expression signatures have been established from bulk sequencing of tumor tissue. Thus, the information of intratumoural heterogeneity could be lost in bulk transcriptomic data, which may influence the performance of these transcriptomic signatures. Emerging achievement from single cell RNA sequencing enable more comprehensively characterize individual cells in tumors [125]. Thirdly, while the information of transcriptional profiles is valuable, they are just part of a big picture. Protein modifications also play an important role in cellular functions, which could influence protein conformation and thereby change its activity [126]. There is no nucleic acid involved in this process.

2.6 The application of nomograms and decision trees in cancer management

2.6.1 Nomogram

In recent years, statistical prediction models have been established in a variety of malignancies. Currently, nomograms have been widely used for cancer prognosis because nomograms combine complex prognostic variables and transform them into a single numerical estimation probability for a given individual. In a nomogram, each variable is presented separately, with a straight line marked with scales and a given magnitude of the variable. Then, add the scores of all variables to get the cumulative score, and match the corresponding outcome in the result scale [127].

The development of a nomogram requires the following steps [128]: Firstly, identifying the patient population where the nomogram will be used and determining the goal of the nomogram. It often refers to a specific event, such as time to recurrence or death. Secondly, the selection of variables is the crucial step that may influence the predictive efficacy of the nomogram. The included prognostic parameters should be clinically significant or have been reported in prior research. Then, a suitable statistical method should be chosen. Logistic regression analysis and Cox proportional hazards are the most used for binary outcome and survival analysis, respectively. Finally, the performance assessment of the predictive model is also needed, including validation, discrimination, and calibration.

Through literature review, there are numerous studies on prognostic nomograms for GC. For example, Yu et al. constructed two prognostic nomograms for OS and cancer-specific survival based on the SEER database using multivariate Cox analysis. The proposed nomograms contain tumor size, location, and stage. The performance of the nomogram was evaluated by the C-index, calibration plot, receiver operating characteristics curve and decision curve analysis. However, this study did not use external verification, and many important clinical factors are missing in the SEER database [129]. A study from Korea developed and externally validated survival-related nomogram for GC based on 7954 patients. The nomogram integrated more valuable information, including age, sex, depth of invasion, location, and the number of metastatic/examined lymph nodes. External validation in two independent cohorts presented excellent discrimination power (C-index, 0.78 and 0.79, respectively) [130]. In addition, there are several published studies that incorporate the risk score of prognostic gene signatures into the nomogram, providing better predictive accuracy than the traditional clinicopathological factors [120, 121, 131-133].

2.6.2 Decision tree

Decision tree (DT) methodology is a non-parametric supervised learning method for establishing classification systems or for constructing prediction models. It is drawn upside down with its root at the top and branches into possible outcomes. Each of those outcomes is linked to internal nodes, which branch off into other results [134]. The commonly-used algorithms include classification and regression trees (CART), ID3s, chi-square automatic interaction detector DTs, and C4.5 and C5.0 DTs. Currently, DTs have been widely applied in the diagnosis of certain diseases

from the patterns of symptoms, or in the identification of the risk stratification of cancer patients based on different prognostic variables.

In 2016, a study from Australia has developed and validated a novel survival prediction model based on a CART algorithm for patients with malignant pleural mesothelioma. The prognostic model identified four risk groups with significant survival differences ($p<0.0001$), which contained five decisive variables. The overall performance of the DT model was rational (C-index= 0.76) and demonstrated a high sensitivity (94.5%) for patients' prognosis [135]. In addition to survival prediction, Takada et al. constructed a DT-based prediction model for predicting axillary lymph node metastasis in patients with primary BC. Fifteen of 24 variables were determined to develop the DT model, including age, BMI and related variables from ultrasound, mammography, physical examination, and pathology. The AUC values of prediction model reached 0.77 for the training cohort and 0.772 for the validation cohort, indicating its high accuracy and strong generalization ability [136]. More recently, several studies combined clinicopathological features and gene signature risk scores to build DT model for risk stratification in different types of cancer. These results indicated that DTs can identify different risk subgroups powerfully and the risk scores from gene signatures often served as the strong determinant [137-139].

However, some limitations of the DT method should be recognized. One of the limitations is that DT-based models can be affected by overfitting and underfitting, particularly when working with small datasets [134]. Another potential problem is that the selection of variables in the decision tree is based on mathematical conditions, which often have different thresholds on different levels of tree. This could make the findings difficult to interpret, as the cutoff values of a certain variable may not be meaningful or validated [140].

2.7 The predictive function of gene signatures for therapeutic response

Cancer therapeutic resistance is the most frequent cause of anti-tumor treatment failure, leading to more cancer-related mortality. Currently, the most common way to determine whether a treatment is effective is based on cancer subtypes and genetic mutations. However, this approach remains imprecise due to the heterogeneity of the tumor. In recent years, high-throughput sequence technology has been used to identify gene expression signatures that can predict the response to treatments, such as chemotherapy and immunotherapy [141-143].

Chemotherapy for various types of cancer is the major factor in reducing recurrence and death. However, due to side effects that occur during chemotherapy, the benefits need to be fully assessed before use on patients [144, 145]. For example, Hess et al. developed a multigene signature for predicting preoperative chemosensitivity in BC patients. The multigene predictor contains 30 optimal probes that were chosen from the differentially expressed genes between pathologic complete response versus residual disease. It showed significantly higher sensitivity (92% vs 61%) than traditional clinical variables [146]. Recently, a multi-cohort study proposed a four gene classifier (GZMB, WARS, SFRP4, and CDX1) for chemotherapy response in resectable GC. The patients who would benefit from chemotherapy, based on the prediction from the gene signature and received adjuvant chemotherapy demonstrated significantly longer 5-year OS compared with the patients who underwent surgery alone ($p=0.0012$) [147].

Over the last decade, although the checkpoint inhibitor-based immunotherapy has achieved clinical success in various types of malignancies, only a minority of patients can obtain a positive response to checkpoint inhibitor therapy. Thus, a predictive biomarker is needed to assess patient benefit. For example, Ribas et al. developed an interferon-γ regulated gene signature in advanced melanoma patients, which demonstrated a significant correlation with both enhanced overall response rates and progression-free survival to anti-PD-1 treatment [148]. Moreover, the T-effector and interferon-γ gene signature that reflects pre-existing immune status, was used in the clinical trial of patients with NSCLC (POPLAR trial). High gene expression signature levels can effectively predict prolonged OS for patients treated with atezolizumab [149]. These results support that the infiltration of immune cells in the tumor microenvironment (TME) possessed great value in predicting immunotherapeutic responses.

More recently, Zeng et al. comprehensively analyzed the TME of GC and identified different TME phenotypes with significant differences in survival [150]. Importantly, the TME score proposed by the study served as a robust biomarker for predicting immunotherapeutic efficacy, which has also been confirmed in other malignancies [151, 152]. Currently, Immune Cell Abundance Identifier (ImmuCellAI), a gene set signature-based algorithm, was established for predicting immunotherapy response by estimating the abundance of specific immune cells from gene expression profiles [153]. This accurate and reliable tool provided an unprecedented opportunity to theoretically predict the patient's immunotherapy efficacy, but it still requires further prospective studies.

2.8 Aims of the present studies

Previously studies about prognostic factors in GC and sarcomas have been published. However, the power of these prognostic tools insufficiently considers tumor heterogeneity. Thus, a novel and precise survival prediction model based on gene expression data should be developed. Identifying more robust prognostic biomarkers will lead to an improved guidance of clinical management and predict prognosis for cancer patients. In addition, gene expression profiling might be helpful in revealing the underlying molecular mechanisms of tumor development.

2.8.1 Establishment of angiogenesis-related gene signatures for prognosis prediction in gastric cancer

Tumor-associated angiogenesis is critical for tumor progression and metastasis. It is necessary to understand the prognostic role of angiogenesis-related genes (ARGs), which might further help to identify novel therapeutic targets. Previous studies have reported that the expression levels of several ARGs in GC correlated with prognosis. However, no such polygenic ARG based risk score was investigated in GC so far.

Therefore, the first study aimed to identify OS-related ARGs and establish an ARG expression signature predicting survival based on public GC datasets. Consensus classification was used to identify the novel molecular subtypes of GC based on ARG expression profiling. Moreover, a

nomogram integrating the new ARG signature risk score and clinicopathological features was developed to quantify individual risk assessment.

2.8.2 Development of an immune-related gene signature predicting survival in sarcoma patients

Sarcomas are a heterogeneous group of rare mesenchymal tumors. Accumulating evidence reveals that the migration of immune cells into these tumors are associated with the clinical outcome of immunotherapy in sarcomas [154, 155]. The current risk stratification system is insufficient to provide precise survival prediction and treatment response.

Therefore, this study explored the association between various immune cells and the prognosis of sarcomas. Different bioinformatics and statistical methods were used to develop a powerful immune-related gene signature for predicting survival outcome and responses to immunotherapy. Moreover, an integrated decision tree and nomogram was established based on the identified gene signature and clinicopathological characteristics to improve risk stratification and quantify the risk assessment of individual patients.

2.8.3 Explore the association between immunosurveillance and organotropism of metastases in colorectal cancer

Approximately 30% of colorectal cancer (CRC) patients present with distant metastases at initial diagnosis, and another 50% of patients will develop metachronous metastases to the liver or the peritoneum. However, no specific biomarkers or gene signatures have been found so far assessing the risk for developing distant metastases. Thus, patients with locally advanced CRCs undergoing surgery were screened and divided into three groups M0 (no distant metastasis), HEP (liver metastasis), and PER (peritoneal carcinomatosis).

Six patients of each group were selected randomly and further assessed by NanoString analysis (the nCounter® PanCancer Progression Panel). Next, gene set enrichment analysis (GSEA) was used to explore potential gene sets related to different metastatic routes.

3. Summary (in Englisch)

3.1 Paper 1:

Angiogenesis-Related Gene Expression Signatures Predicting Prognosis in Gastric Cancer Patients

Tumor angiogenesis plays a crucial role in the occurrence and development of gastric cancer (GC). Thus, it is critical to identify an angiogenesis-related gene (ARG) signature that serve as prognostic biomarkers for GC patients.

The ARG set was downloaded from hallmark-angiogenesis in “Gene Set Enrichment Analysis” (GSEA), which contains 36 genes involved in tumor angiogenesis. The expression data of ARGs and clinicopathological features of 372 GC patients were obtained from The Cancer Genome Atlas (TCGA). Consensus clustering was used to identify angiogenesis-related subtypes (cluster 1 and 2). Least absolute shrinkage and selection operator (LASSO) Cox regression analyses were performed in the training cohort (TCGA) to screen out the overall survival-related ARGs and establish prognostic gene signatures, respectively. The ARG-risk score was calculated based on the gene expression value and the corresponding coefficients in the ARG signature. Then patients were divided into high- and low-risk score groups for survival analysis. The Asian Cancer Research Group (ACRG) (n = 300) was used as an external validation cohort. The ARG signature and relevant clinical features were used to establish the prognostic nomogram. GSEA was performed to explore the potential signaling pathways related to ARG-risk score.

Consensus clustering revealed that cluster 2 patients exhibited a more advanced clinical stage and had worse survival outcomes than patients in cluster 1. An optimal set of 10 ARGs was identified and used to construct the prognostic ARG signature. The signature showed prognostic significant relevance in both training and validation cohorts. In multivariate analysis, the ARG-risk score proved to be an independent prognostic indicator for disease free and overall survival. The predictive ability of the nomogram containing the risk score and clinicopathological information was superior to TNM staging. In the high-risk score group, several cancer and metastasis-related signaling pathways were significantly overexpressed compared with the low-risk score group.

In conclusion, our 10-ARG signature is closely associated with survival outcome in GC patients and might be useful for a more accurate risk stratification.

3.2 Paper 2:

A novel immune-related gene signature predicting survival in sarcoma patients

Sarcoma patients exhibit significant heterogeneity in survival outcome. The current risk stratification system is insufficient to provide precise survival prediction and treatment response. Therefore, an effective model is needed to predict prognosis and guide treatment.

This study analyzed the gene expression and outcome of 980 sarcoma patients from seven public datasets, including six microarray datasets from Gene Expression Omnibus (GEO) and one RNA-Seq dataset from The Cancer Genome Atlas (TCGA). The abundance of immune cells and the response to immunotherapy was calculated in the ImmuCellAI database. Immune-related genes (IRG) were screened through weighted gene co-expression network analysis (WGCNA). Least absolute shrinkage and selection operator (LASSO) Cox regression was further used to develop

the prognostic IRG signature. We divided patients into high and low-risk groups based on the risk score of the IRG signature and compared survival between groups. Moreover, we developed a nomogram and a decision tree integrating the IRG risk score and clinicopathological parameters.

The identified IRG signature incorporated 14 genes (TRIM21, TNF, CRIP1, FCER1A, SLC25A20, ZNFX1, DHX58, TNFSF15, TAPBPL, CMA1, APOL2 and GBP2) and was significantly associated with survival outcome in the training and six independent validation cohorts. Multivariate survival analyses revealed that the 14-IRG signature served as an independent risk factor for disease free and overall survival. Moreover, the IRG signature acted as a potential indicator for immunotherapy. The nomogram based on the IRG signature risk score outperformed presented traditional clinicopathological features in survival prediction. In addition, the decision tree discriminated risk subgroups powerfully.

In summary, the proposed IRG signature is a reliable predictive tool to predict outcome and treatment response in sarcoma patients.

3.3 Paper 3:

The association of immunosurveillance and distant metastases in colorectal cancer

Colorectal cancer (CRC) is one of the most common malignancies worldwide with still an increasing incidence, but little is known about the underlying mechanism of distant metastases. The purpose of this study was to evaluate the association between immunosurveillance and organotropism of metastases in CRC.

CRC patients undergoing surgery with a 5-year follow-up at the Ludwig-Maximilian University Hospital Munich were selected and divided into three groups: M0 (no distant metastases), HEP (liver metastases) and PER (peritoneal carcinomatosis). Each group randomly selects six patients for a NanoString analysis, which includes 770 genes from 13 cancer-associated canonical pathways. Then, gene set enrichment analysis (GSEA) was further performed based on seven main cancer-associated databases.

Comparing HEP vs. M0, the gene set associated with the Toll-like receptor (TLR) cascade was highly enriched in HEP group. HSP90B1, MAPKAPK3, PPP2CB, PPP2R1A had the greatest influence on the core enrichment. In the M0 group, the immunologic signature pathway GSE6875_TCONV_VS_FOXP3_KO_TREG_DN was significantly overexpressed compared with HEP group. RB1, TMEM 100, CFP, ZKSCAN5, DDX50 contributed the most to enrichment analysis.

In this study, differential expressed immune related gene sets were investigated between CRC with either hepatic or peritoneal metastases, indicating that the occurrence of liver metastases might be related to TLR-associated pathways and immunosurveillance mediated by FOXP3.

4. Zusammenfassung (deutsch)

4.1 Paper 1:

Angiogenesis-Related Gene Expression Signatures Predicting Prognosis in Gastric Cancer Patients

Die Tumorangiogenese spielt eine entscheidende Rolle bei der Entstehung und Entwicklung von Magenkrebs (GC). Daher erscheint es sinnvoll, die prognostische Wertigkeit einer Angiogenese-bezogene Gensignatur (ARG) bei GC-Patienten zu untersuchen.

Von der "Gene Set Enrichment Analyse" (GSEA) der „Hallmark-Angiogenese“ wurde hierzu ein Set aus 36 Genen identifiziert, die an der Tumorangiogenese beteiligt sind. Die Expressionsdaten der ARGs und die klinisch-pathologischen Merkmale von 372 GC-Patienten wurden anschließend von der öffentlichen Datenbank „The Cancer Genome Atlas“ (TCGA) entnommen. Zur Identifizierung der mit der Angiogenese zusammenhängenden Subtypen (Cluster 1 und 2) wurde ein Konsens-Clustering durchgeführt. In der Trainingskohorte (TCGA) erfolgten sodann Cox-Regressionsanalysen mit dem LASSO-Verfahren (Least Absolute Shrinkage and Selection Operator), um die mit dem Gesamtüberleben verbundenen ARGs herauszufiltern bzw. prognostische Gensignaturen zu erstellen. Basierend auf diesen Genexpressionswerten und der entsprechenden Koeffizienten konnte eine prognostisch relevante ARG-Signatur und konsekutiv ein ARG-Risk-Score berechnet werden. Anschließend wurden die Patienten für die Überlebensanalyse in Gruppen mit hohem und niedrigem Risikoscore eingeteilt. Die öffentliche Datenbank „Asian Cancer Research Group“ (ACRG) (n = 300) wurde als externe Validierungskohorte verwendet. Die ARG-Signatur und relevante klinische Merkmale wurden zur Erstellung eines prognostischen Nomogramms verwendet. Mit Hilfe von GSEA wurden die potenziellen Signalwege untersucht, die mit dem ARG-Risikoscore in Verbindung stehen.

Das Konsens-Clustering ergab, dass Patienten in Cluster 2 (hoher ARG-Risk-Score) ein fortgeschritteneres klinisches Stadium und ein schlechteres Überleben als Patienten in Cluster 1 aufwiesen. Die Analysen ergaben, dass die prognostische ARG-Signatur aus 10 ARGs besteht. Die Signatur zeigte sowohl in den Trainings- als auch in den Validierungskohorten eine signifikante prognostische Relevanz. In der multivariaten Analyse erwies sich der ARG-Risiko-Score als unabhängiger prognostischer Indikator für das Gesamt- und krankheitsfreie Überleben. Die Vorhersagekraft des Nomogramms, das den Risikoscore und klinisch-pathologische Informationen enthält, war dem herkömmlichen TNM-Staging überlegen. In der Gruppe mit hohem Risikoscore waren mehrere krebs- und metastasenbezogene Signalwege im Vergleich zur Gruppe mit niedrigem Risikoscore deutlich überexprimiert.

Zusammenfassend lässt sich sagen, dass unsere 10-ARG-Signatur eng mit dem Überleben von GC-Patienten assoziiert ist und für eine genauere Risikostratifizierung nützlich sein könnte.

4.2 Paper 2:

A novel immune-related gene signature predicting survival in sarcoma patients

Sarkom-Patienten weisen eine erhebliche Heterogenität im Überlebensverlauf auf. Das derzeitige System zur Risikostratifizierung erscheint nicht ausreichend zu sein, um eine präzise Vorhersage

des Überlebens und des Ansprechens auf die Behandlung zu ermöglichen. Daher ist ein wirksames Modell erforderlich, um die Prognose vorherzusagen und die Behandlung zu steuern.

In dieser Studie wurden daher die Genexpression und das Ergebnis von 980 Sarkom-Patienten aus sieben öffentlichen Datensätzen analysiert, darunter sechs Microarray-Datensätze aus dem „Gene Expression Omnibus“ (GEO) und ein RNA-Seq-Datensatz aus dem „The Cancer Genome Atlas“ (TCGA). Die Häufigkeit von Immunzellen und das Ansprechen auf eine Immuntherapie wurden in der ImmuCellAI-Datenbank berechnet. Immunbezogene Gene (IRG) wurden mit Hilfe der gewichteten Gen-Koexpressionsnetzwerk-Analyse (WGCNA) untersucht. Zur Entwicklung der prognostischen IRG-Signatur und einem darauf basierenden Risikoscore wurde die LASSO-Cox-Regression Analyse (Least Absolute Shrinkage and Selection Operator) verwendet. Wir teilten die Patienten anhand des Risikoscores der IRG-Signatur in Hoch- und Niedrigrisikogruppen ein und verglichen das Überleben zwischen den Gruppen. Darüber hinaus entwickelten wir ein Nomogramm und einen Entscheidungsbaum, die den IRG-Risikoscore und klinisch-pathologische Parameter integrierten.

Die identifizierte IRG-Signatur umfasste 14 Gene (TRIM21, TNF, CRIP1, FCER1A, SLC25A20, ZNFX1, DHX58, TNFSF15, TAPBPL, CMA1, APOL2 und GBP2), die in der Trainings- und sechs unabhängigen Validierungskohorten signifikant mit dem Gesamtüberleben assoziiert waren. Multivariate Überlebensanalysen zeigten, dass die 14-IRG-Signatur einen unabhängigen Risikofaktor für das Gesamt- und krankheitsfreie Überleben darstellt. Darüber hinaus fungierte die IRG-Signatur als potenzieller Indikator für eine Immuntherapie. Das Nomogramm, das auf dem IRG-Signatur-Risikoscore basierte, übertraf die traditionellen klinisch-pathologischen Merkmale bei der Überlebensvorhersage. Der Entscheidungsbaum ermöglichte außerdem eine gute Unterscheidung von Risikogruppen.

Zusammenfassend lässt sich sagen, dass die vorgeschlagene 14-IRG-Signatur ein zuverlässiges Instrument zur Vorhersage des Ergebnisses und des Ansprechens auf die Behandlung bei Sarkom-Patienten ist.

4.3 Paper 3:

The association of immunosurveillance and distant metastases in colorectal cancer

Darmkrebs (CRC) ist eine der häufigsten bösartigen Erkrankungen mit weltweit steigender Inzidenz, aber es ist nur wenig über den zugrunde liegenden Mechanismus der Fernmetastasierung bekannt. Ziel dieser Studie war es, den Zusammenhang zwischen der Immunüberwachung und dem Organotropismus von Metastasen bei Darmkrebs zu untersuchen.

CRC-Patienten, die am Klinikum der Ludwig-Maximilians-Universität München operiert wurden, wurden in drei Gruppen eingeteilt: M0 (keine Fernmetastasen), HEP (Lebermetastasen) und PER (Peritonealkarzinomatose). Aus jeder Gruppe wurden sechs Patienten nach dem Zufallsprinzip für eine NanoString-Analyse ausgewählt, die 770 Gene aus 13 krebsassoziierten kanonischen Stoffwechselwegen umfasst. Anschließend wurde eine "Gene Set Enrichment Analyse" (GSEA) auf der Grundlage von sieben wichtigen krebsassoziierten Datenbanken durchgeführt.

Im Vergleich zwischen HEP und M0 war der mit der Toll-like-Rezeptor (TLR)-Kaskade assoziierte Gensatz in der HEP-Gruppe stark angereichert. HSP90B1, MAPKAPK3, PPP2CB und PPP2R1A hatten den stärksten Einfluss auf die signifikante Anreicherung. In der M0-Gruppe war der immunologische Signaturpfad GSE6875_TCONV_VS_FOXP3_KO_TREG_DN im Vergleich zur

HEP-Gruppe signifikant überexprimiert. RB1, TMEM 100, CFP, ZKSCAN5 und DDX50 trugen am meisten zur Anreicherungsanalyse bei.

In dieser Studie wurden differentiell exprimierte immunbezogene Gensätze von CRC mit unterschiedlichen Metastasierungsmustern untersucht, was darauf hindeuten könnte, dass das Auftreten von Lebermetastasen mit TLR-assoziierten Stoffwechselwegen und der durch FOXP3 vermittelten Immunüberwachung zusammenhängen könnte.

5. Paper I

Cancers (Basel). 2020 Dec 8;12(12):3685. doi: 10.3390/cancers12123685. PMID: 33302481

Angiogenesis-Related Gene Expression Signatures Predicting Prognosis in Gastric Cancer Patients.

Ren, H.; Zhu, J.; Yu, H.; Bazhin, A.V.; Westphalen, C.B.; Renz, B.W.; Jacob, S.N.; Lampert, C.; Werner, J.; Angele, M.K.; Bös

6. Paper II

Mol Ther Oncolytics. 2021 Dec 9;24:114-126. doi: 10.1016/j.omto.2021.12.007. PMID: 35024438

A novel immune-related gene signature predicting survival in sarcoma patients

Ren H, Bazhin AV, Pretzsch E, Jacob S, Yu H, Zhu J, Albertsmeier M, Lindner LH, Knösel T, Werner J, Angele MK, Bösch F

7. Paper III

J Cancer Res Clin Oncol. 2021 Nov;147(11):3333-3341. doi: 10.1007/s00432-021-03753-w. Epub 2021 Sep 2. PMID: 34476575

The association of immunosurveillance and distant metastases in colorectal cancer.

Jacob S, Jurinovic V, Lampert C, Pretzsch E, Kumbrink J, Neumann J, **Haoyu R**, Renz BW, Kirchner T, Guba MO, Werner J, Angele MK, Bösch F.

References

[1] H. Sung *et al.*, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA Cancer J Clin,* vol. 71, no. 3, pp. 209-249, May 2021, doi: 10.3322/caac.21660.

[2] V. Ficarra, A. Galfano, M. Mancini, G. Martignoni, and W. Artibani, "TNM staging system for renal-cell carcinoma: current status and future perspectives," *Lancet Oncol,* vol. 8, no. 6, pp. 554-8, Jun 2007, doi: 10.1016/S1470-2045(07)70173-0.

[3] E. E. Zervos *et al.*, "Stage does not predict survival after resection of hilar cholangiocarcinomas promoting an aggressive operative approach," *Am J Surg,* vol. 190, no. 5, pp. 810-5, Nov 2005, doi: 10.1016/j.amjsurg.2005.07.025.

[4] M. M. Al-Hawary *et al.*, "Pancreatic ductal adenocarcinoma radiology reporting template: consensus statement of the society of abdominal radiology and the american pancreatic association," *Gastroenterology,* vol. 146, no. 1, pp. 291-304 e1, Jan 2014, doi: 10.1053/j.gastro.2013.11.004.

[5] L. Jouffret, O. Turrini, J. Ewald, V. Moutardier, J. L. Iovanna, and J. R. Delpero, "Long-term survivors after pancreatectomy for cancer: the TNM classification is outdated," *ANZ J Surg,* vol. 85, no. 11, pp. 860-4, Nov 2015, doi: 10.1111/ans.13277.

[6] P. Zhang *et al.*, "(18)F-fluorodeoxyglucose positron emission computed tomography for monitoring tumor response in esophageal carcinoma treated with concurrent chemoradiotherapy," *Oncol Lett,* vol. 15, no. 2, pp. 1845-1852, Feb 2018, doi: 10.3892/ol.2017.7528.

[7] F. L. Greene and L. H. Sobin, "The staging of cancer: a retrospective and prospective appraisal," *CA Cancer J Clin,* vol. 58, no. 3, pp. 180-90, May-Jun 2008, doi: 10.3322/CA.2008.0001.

[8] M. Gospodarowicz and B. O'Sullivan, "Prognostic factors in cancer," *Semin Surg Oncol,* vol. 21, no. 1, pp. 13-8, 2003, doi: 10.1002/ssu.10016.

[9] R. W. Tsang, J. D. Brierley, W. J. Simpson, T. Panzarella, M. K. Gospodarowicz, and S. B. Sutcliffe, "The effects of surgery, radioiodine, and external radiation therapy on the clinical outcome of patients with differentiated thyroid carcinoma," *Cancer,* vol. 82, no. 2, pp. 375-88, Jan 15 1998. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/9445196.

[10] K. Derwinger and B. Gustavsson, "A study of aspects on gender and prognosis in synchronous colorectal cancer," *Clin Med Insights Oncol,* vol. 5, pp. 259-64, 2011, doi: 10.4137/CMO.S7871.

[11] R. T. Chlebowski *et al.*, "Ethnicity and breast cancer: factors influencing differences in incidence and outcome," *J Natl Cancer Inst,* vol. 97, no. 6, pp. 439-48, Mar 16 2005, doi: 10.1093/jnci/dji064.

[12] S. G. Craig *et al.*, "Immune status is prognostic for poor survival in colorectal cancer patients and is associated with tumour hypoxia," *Br J Cancer,* vol. 123, no. 8, pp. 1280-1288, Oct 2020, doi: 10.1038/s41416-020-0985-5.

[13] S. E. Oh *et al.*, "Prognostic significance of perioperative nutritional parameters in patients with gastric cancer," *Clin Nutr,* vol. 38, no. 2, pp. 870-876, Apr 2019, doi: 10.1016/j.clnu.2018.02.015.

[14] M. L. Janssen-Heijnen, S. Smulders, V. E. Lemmens, F. W. Smeenk, H. J. van Geffen, and J. W. Coebergh, "Effect of comorbidity on the treatment and prognosis of elderly patients with non-small cell lung cancer," *Thorax,* vol. 59, no. 7, pp. 602-7, Jul 2004, doi: 10.1136/thx.2003.018044.

[15] B. D. Beutler *et al.*, "Sociodemographic Characteristics as Predictors of Outcomes in Hepatocellular Carcinoma: A Retrospective Cohort Study," *Cancer Control,* vol. 27, no. 1, p. 1073274820956615, Jan-Dec 2020, doi: 10.1177/1073274820956615.

[16] H. H. Hartgrink and C. J. van de Velde, "Status of extended lymph node dissection: locoregional control is the only way to survive gastric cancer," *J Surg Oncol,* vol. 90, no. 3, pp. 153-65, Jun 1 2005, doi: 10.1002/jso.20222.

[17] A. R. Novotny, C. Schuhmacher, R. Busch, M. W. Kattan, M. F. Brennan, and J. R. Siewert, "Predicting individual survival after gastric cancer resection: validation of a U.S.-derived nomogram at a single high-volume center in Europe," *Ann Surg,* vol. 243, no. 1, pp. 74-81, Jan 2006, doi: 10.1097/01.sla.0000194088.81126.85.

[18] J. R. Siewert, K. Bottcher, H. J. Stein, and J. D. Roder, "Relevant prognostic factors in gastric cancer: ten-year results of the German Gastric Cancer Study," *Ann Surg,* vol. 228, no. 4, pp. 449-61, Oct 1998, doi: 10.1097/00000658-199810000-00002.

[19] J. P. Kim, Y. W. Kim, H. K. Yang, and D. Y. Noh, "Significant prognostic factors by multivariate analysis of 3926 gastric cancer patients," *World J Surg,* vol. 18, no. 6, pp. 872-7; discussion 877-8, Nov-Dec 1994, doi: 10.1007/BF00299091.

[20] J. P. Kim, J. H. Lee, S. J. Kim, H. J. Yu, and H. K. Yang, "Clinicopathologic characteristics and prognostic factors in 10 783 patients with gastric cancer," *Gastric Cancer,* vol. 1, no. 2, pp. 125-133, Mar 1998, doi: 10.1007/s101200050006.

[21] J. B. Koea, M. S. Karpeh, and M. F. Brennan, "Gastric cancer in young patients: demographic, clinicopathological, and prognostic factors in 92 patients," *Ann Surg Oncol,* vol. 7, no. 5, pp. 346-51, Jun 2000, doi: 10.1007/s10434-000-0346-9.

[22] P. Jimenez Fonseca *et al.*, "Lauren subtypes of advanced gastric cancer influence survival and response to chemotherapy: real-world data from the AGAMENON National Cancer Registry," *Br J Cancer,* vol. 117, no. 6, pp. 775-782, Sep 5 2017, doi: 10.1038/bjc.2017.245.

[23] B. Zhao *et al.*, "Assessment of the 8th edition of TNM staging system for gastric cancer: the results from the SEER and a single-institution database," *Future Oncol,* vol. 14, no. 29, pp. 3023-3035, Dec 2018, doi: 10.2217/fon-2018-0299.

[24] L. Marano *et al.*, "Comparison between 7th and 8th edition of AJCC TNM staging system for gastric cancer: old problems and new perspectives," *Transl Gastroenterol Hepatol,* vol. 4, p. 22, 2019, doi: 10.21037/tgh.2019.03.09.

[25] W. Yasui, N. Oue, P. P. Aung, S. Matsumura, M. Shutoh, and H. Nakayama, "Molecular-pathological prognostic factors of gastric cancer: a review," *Gastric Cancer,* vol. 8, no. 2, pp. 86-94, 2005, doi: 10.1007/s10120-005-0320-0.

[26] C. Chen *et al.*, "Prognostic role of human epidermal growth factor receptor in gastric cancer: a systematic review and meta-analysis," *Arch Med Res,* vol. 44, no. 5, pp. 380-9, Jul 2013, doi: 10.1016/j.arcmed.2013.07.001.

[27] G. Galizia *et al.*, "Epidermal growth factor receptor (EGFR) expression is associated with a worse prognosis in gastric cancer patients undergoing curative surgery," *World J Surg,* vol. 31, no. 7, pp. 1458-68, Jul 2007, doi: 10.1007/s00268-007-9016-4.

[28] E. Lieto *et al.*, "Expression of vascular endothelial growth factor (VEGF) and epidermal growth factor receptor (EGFR) is an independent prognostic indicator of worse outcome in gastric cancer patients," *Ann Surg Oncol,* vol. 15, no. 1, pp. 69-79, Jan 2008, doi: 10.1245/s10434-007-9596-0.

[29] H. Saito *et al.*, "The expression of transforming growth factor-beta1 is significantly correlated with the expression of vascular endothelial growth factor and poor prognosis of patients with advanced gastric carcinoma," *Cancer,* vol. 86, no. 8, pp. 1455-62, Oct 15 1999, doi: 10.1002/(sici)1097-0142(19991015)86:8<1455::aid-cncr11>3.0.co;2-l.

[30] F. J. Kubben *et al.*, "Matrix metalloproteinase-2 is a consistent prognostic factor in gastric cancer," *Br J Cancer,* vol. 94, no. 7, pp. 1035-40, Apr 10 2006, doi: 10.1038/sj.bjc.6603041.

[31] S. Soleyman-Jahi, S. Nedjat, A. Abdirad, N. Hoorshad, R. Heidari, and K. Zendehdel, "Prognostic significance of matrix metalloproteinase-7 in gastric cancer survival: a meta-analysis," *PLoS One,* vol. 10, no. 4, p. e0122316, 2014, doi: 10.1371/journal.pone.0122316.

[32] Q. W. Zhang *et al.*, "Matrix metalloproteinase-9 as a prognostic factor in gastric cancer: a meta-analysis," *Asian Pac J Cancer Prev,* vol. 13, no. 6, pp. 2903-8, 2012, doi: 10.7314/apjcp.2012.13.6.2903.

[33] M. G. Choi *et al.*, "Expression levels of cyclin G2, but not cyclin E, correlate with gastric cancer progression," *J Surg Res,* vol. 157, no. 2, pp. 168-74, Dec 2009, doi: 10.1016/j.jss.2008.06.020.

[34] Y. Takano, Y. Kato, M. Masuda, Y. Ohshima, and I. Okayasu, "Cyclin D2, but not cyclin D1, overexpression closely correlates with gastric cancer progression and prognosis," *J Pathol,* vol. 189, no. 2, pp. 194-200, Oct 1999, doi: 10.1002/(SICI)1096-9896(199910)189:2<194::AID-PATH426>3.0.CO;2-P.

[35] J. D. Harrison, "The gastrointestinal absorption of the actinide elements," *Sci Total Environ,* vol. 100 Spec No, pp. 43-60, Mar 1991, doi: 10.1016/0048-9697(91)90373-m.

[36] M. Yildirim, V. Kaya, O. Demirpence, S. Gunduz, and H. Bozcuk, "Prognostic significance of p53 in gastric cancer: a meta- analysis," *Asian Pac J Cancer Prev,* vol. 16, no. 1, pp. 327-32, 2015, doi: 10.7314/apjcp.2015.16.1.327.

[37] T. Sasaki *et al.*, "AKT activation and telomerase reverse transcriptase expression are concurrently associated with prognosis of gastric cancer," *Pathobiology,* vol. 81, no. 1, pp. 36-41, 2014, doi: 10.1159/000351721.

[38] T. Kondo *et al.*, "Expression of POT1 is associated with tumor stage and telomere length in gastric carcinoma," *Cancer Res,* vol. 64, no. 2, pp. 523-9, Jan 15 2004, doi: 10.1158/0008-5472.can-03-1196.

[39] X. Xing *et al.*, "The prognostic value of E-cadherin in gastric cancer: a meta-analysis," *Int J Cancer,* vol. 132, no. 11, pp. 2589-96, Jun 1 2013, doi: 10.1002/ijc.27947.

[40] Y. Chen, Z. Fu, S. Xu, Y. Xu, and P. Xu, "The prognostic value of CD44 expression in gastric cancer: a meta-analysis," *Biomed Pharmacother,* vol. 68, no. 6, pp. 693-7, Jul 2014, doi: 10.1016/j.biopha.2014.08.001.

[41] J. W. Xie *et al.*, "Evaluation of the prognostic value and functional roles of CD44v6 in gastric cancer," *J Cancer Res Clin Oncol,* vol. 141, no. 10, pp. 1809-17, Oct 2015, doi: 10.1007/s00432-015-1964-8.

[42] M. Ratti, A. Lampis, J. C. Hahne, R. Passalacqua, and N. Valeri, "Microsatellite instability in gastric cancer: molecular bases, clinical perspectives, and new treatment approaches," *Cell Mol Life Sci,* vol. 75, no. 22, pp. 4151-4162, Nov 2018, doi: 10.1007/s00018-018-2906-9.

[43] W. L. Fang *et al.*, "Microsatellite instability is associated with a better prognosis for gastric cancer patients after curative surgery," *World J Surg,* vol. 36, no. 9, pp. 2131-8, Sep 2012, doi: 10.1007/s00268-012-1652-7.

[44] Y. Y. Choi *et al.*, "Is microsatellite instability a prognostic marker in gastric cancer? A systematic review with meta-analysis," *J Surg Oncol,* vol. 110, no. 2, pp. 129-35, Aug 2014, doi: 10.1002/jso.23618.

[45] H. S. Lee *et al.*, "Distinct clinical features and outcomes of gastric cancers with microsatellite instability," *Mod Pathol,* vol. 15, no. 6, pp. 632-40, Jun 2002, doi: 10.1038/modpathol.3880578.

[46] E. C. Smyth *et al.*, "Mismatch Repair Deficiency, Microsatellite Instability, and Survival: An Exploratory Analysis of the Medical Research Council Adjuvant Gastric Infusional Chemotherapy (MAGIC) Trial," *JAMA Oncol,* vol. 3, no. 9, pp. 1197-1203, Sep 1 2017, doi: 10.1001/jamaoncol.2016.6762.

[47] D. G. Fu, "Epigenetic alterations in gastric cancer (Review)," *Mol Med Rep,* vol. 12, no. 3, pp. 3223-3230, Sep 2015, doi: 10.3892/mmr.2015.3816.

[48] Y. Peng, Q. Wu, L. Wang, H. Wang, and F. Yin, "A DNA methylation signature to improve survival prediction of gastric cancer," *Clin Epigenetics,* vol. 12, no. 1, p. 15, Jan 20 2020, doi: 10.1186/s13148-020-0807-x.

[49] C. Li *et al.*, "A four-DNA methylation signature as a novel prognostic biomarker for survival of patients with gastric cancer," *Cancer Cell Int,* vol. 20, p. 88, 2020, doi: 10.1186/s12935-020-1156-8.

[50] X. Ma, H. Chen, G. Wang, L. Li, and K. Tao, "DNA methylation profiling to predict overall survival risk in gastric cancer: development and validation of a nomogram to optimize clinical management," *J Cancer,* vol. 11, no. 15, pp. 4352-4365, 2020, doi: 10.7150/jca.44436.

[51] B. A. Nacev *et al.*, "The epigenomics of sarcoma," *Nat Rev Cancer,* vol. 20, no. 10, pp. 608-623, Oct 2020, doi: 10.1038/s41568-020-0288-4.

[52] C. A. Stiller, A. W. Craft, I. Corazziari, and E. W. Group, "Survival of children with bone sarcoma in Europe since 1978: results from the EUROCARE study," *Eur J Cancer,* vol. 37, no. 6, pp. 760-6, Apr 2001, doi: 10.1016/s0959-8049(01)00004-1.

[53] M. A. Clark, C. Fisher, I. Judson, and J. M. Thomas, "Soft-tissue sarcomas in adults," *N Engl J Med,* vol. 353, no. 7, pp. 701-11, Aug 18 2005, doi: 10.1056/NEJMra041866.

[54] K. Maretty-Nielsen, "Prognostic factors in soft tissue sarcoma," *Dan Med J,* vol. 61, no. 11, p. B4957, Nov 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/25370967.

[55] J. LeVay *et al.*, "Outcome and prognostic factors in soft tissue sarcoma in the adult," *Int J Radiat Oncol Biol Phys,* vol. 27, no. 5, pp. 1091-9, Dec 1 1993, doi: 10.1016/0360-3016(93)90529-5.

[56] E. A. Levine, "Prognostic factors in soft tissue sarcoma," *Semin Surg Oncol,* vol. 17, no. 1, pp. 23-32, Jul-Aug 1999, doi: 10.1002/(sici)1098-2388(199907/08)17:1<23::aid-ssu4>3.0.co;2-r.

[57] P. D. Stefanovski *et al.*, "Prognostic factors in soft tissue sarcomas: a study of 395 patients," *Eur J Surg Oncol,* vol. 28, no. 2, pp. 153-64, Mar 2002, doi: 10.1053/ejso.2001.1242.

[58] M. Amin, S. Edge, and F. Greene, *AJCC Cancer Staging Manual*. Springer International Publishing (in English), 2017.

[59] E. Youssef *et al.*, "Long-term outcome of combined modality therapy in retroperitoneal and deep-trunk soft-tissue sarcoma: analysis of prognostic factors," *Int J Radiat Oncol Biol Phys,* vol. 54, no. 2, pp. 514-9, Oct 1 2002, doi: 10.1016/s0360-3016(02)02942-5.

[60] G. Xu *et al.*, "Homogenous and Heterogenous Prognostic Factors for Patients with Bone Sarcoma," *Orthop Surg,* vol. 13, no. 1, pp. 134-144, Feb 2021, doi: 10.1111/os.12851.

[61] S. S. Bielack *et al.*, "Prognostic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1,702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols," *J Clin Oncol,* vol. 20, no. 3, pp. 776-90, Feb 1 2002, doi: 10.1200/JCO.2002.20.3.776.

[62] T. Ozaki *et al.*, "Osteosarcoma of the spine: experience of the Cooperative Osteosarcoma Study Group," *Cancer,* vol. 94, no. 4, pp. 1069-77, Feb 15 2002. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/11920477.

[63] R. L. Huuhtanen *et al.*, "Comparison of the Ki-67 score and S-phase fraction as prognostic variables in soft-tissue sarcoma," *Br J Cancer,* vol. 79, no. 5-6, pp. 945-51, Feb 1999, doi: 10.1038/sj.bjc.6690151.

[64] P. F. Choong *et al.*, "Expression of proliferating cell nuclear antigen (PCNA) and Ki-67 in soft tissue sarcoma. Is prognostic significance histotype-specific?," *APMIS,* vol. 103, no. 11, pp. 797-805, Nov 1995, doi: 10.1111/j.1699-0463.1995.tb01437.x.

[65] B. T. Skytting, H. C. Bauer, R. Perfekt, G. Nilsson, and O. Larsson, "Ki-67 is strongly prognostic in synovial sarcoma: analysis based on 86 patients from the Scandinavian Sarcoma group register," *Br J Cancer,* vol. 80, no. 11, pp. 1809-14, Aug 1999, doi: 10.1038/sj.bjc.6690602.

[66] J. M. Lopes *et al.*, "Synovial sarcoma. Evaluation of prognosis with emphasis on the study of DNA ploidy and proliferation (PCNA and Ki-67) markers," *Anal Cell Pathol,* vol. 16, no. 1, pp. 45-62, 1998, doi: 10.1155/1998/545906.

[67] M. Zeng, J. Zhou, L. Wen, Y. Zhu, Y. Luo, and W. Wang, "The relationship between the expression of Ki-67 and the prognosis of osteosarcoma," *BMC Cancer,* vol. 21, no. 1, p. 210, Mar 1 2021, doi: 10.1186/s12885-021-07880-y.

[68] X. Wang *et al.*, "The prognostic value of PCNA expression in patients with osteosarcoma: A meta-analysis of 16 studies," *Medicine (Baltimore),* vol. 96, no. 41, p. e8254, Oct 2017, doi: 10.1097/MD.0000000000008254.

[69] F. Collin, A. Chassevent, F. Bonichon, G. Bertrand, P. Terrier, and J. M. Coindra, "Flow cytometric DNA content analysis of 185 soft tissue neoplasms indicates that S-phase fraction is a prognostic factor for sarcomas. French Federation of Cancer Centers (FNCLCC) Sarcoma Group," *Cancer,* vol. 79, no. 12, pp. 2371-9, Jun 15 1997, doi: 10.1002/(sici)1097-0142(19970615)79:12<2371::aid-cncr11>3.0.co;2-o.

[70] L. C. Wijnaendts *et al.*, "Prognostic importance of DNA flow cytometric variables in rhabdomyosarcomas," *J Clin Pathol,* vol. 46, no. 10, pp. 948-52, Oct 1993, doi: 10.1136/jcp.46.10.948.

[71] P. Gustafson *et al.*, "Flow cytometric S-phase fraction in soft-tissue sarcoma: prognostic importance analysed in 160 patients," *Br J Cancer,* vol. 75, no. 1, pp. 94-100, 1997, doi: 10.1038/bjc.1997.15.

[72] A. Kawai *et al.*, "Nuclear immunoreaction of p53 protein in soft tissue sarcomas. A possible prognostic factor," *Cancer,* vol. 73, no. 10, pp. 2499-505, May 15 1994, doi: 10.1002/1097-0142(19940515)73:10<2499::aid-cncr2820731008>3.0.co;2-g.

[73] H. Medina-Franco *et al.*, "Expression of p53 and proliferation index as prognostic factors in gastrointestinal sarcomas," *Ann Surg Oncol,* vol. 10, no. 2, pp. 190-5, Mar 2003, doi: 10.1245/aso.2003.03.033.

[74] A. Abudu *et al.*, "Overexpression of p53 protein in primary Ewing's sarcoma of bone: relationship to tumour stage, response and prognosis," *Br J Cancer,* vol. 79, no. 7-8, pp. 1185-9, Mar 1999, doi: 10.1038/sj.bjc.6690190.

[75] R. Schneider-Stock, D. Onnasch, C. Haeckel, W. Mellin, D. S. Franke, and A. Roessner, "Prognostic significance of p53 gene mutations and p53 protein expression in synovial sarcomas," *Virchows Arch,* vol. 435, no. 4, pp. 407-12, Oct 1999, doi: 10.1007/s004280050418.

[76] D. Yao, G. H. Cai, J. Chen, R. Ling, S. X. Wu, and Y. P. Li, "Prognostic value of p53 alterations in human osteosarcoma: a meta analysis," *Int J Clin Exp Pathol,* vol. 7, no. 10, pp. 6725-33, 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/25400752.

[77] C. Cordon-Cardo *et al.*, "Molecular abnormalities of mdm2 and p53 genes in adult soft tissue sarcomas," *Cancer Res,* vol. 54, no. 3, pp. 794-9, Feb 1 1994. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/8306343.

[78] P. Wurl *et al.*, "High prognostic significance of Mdm2/p53 co-overexpression in soft tissue sarcomas of the extremities," *Oncogene,* vol. 16, no. 9, pp. 1183-5, Mar 5 1998, doi: 10.1038/sj.onc.1201646.

[79] J. Shen *et al.*, "Prognostic significance of nuclear accumulation of c-myc and mdm2 proteins in synovial sarcoma of the extremities," *Oncology,* vol. 58, no. 3, pp. 253-60, Apr 2000, doi: 10.1159/000012109.

[80] G. Gamberi *et al.*, "C-myc and c-fos in human osteosarcoma: prognostic value of mRNA and protein expression," *Oncology,* vol. 55, no. 6, pp. 556-63, Nov-Dec 1998, doi: 10.1159/000011912.

[81] M. R. Sollazzo *et al.*, "Increased c-myc oncogene expression in Ewing's sarcoma: correlation with Ki67 proliferation index," *Tumori,* vol. 85, no. 3, pp. 167-73, May-Jun 1999. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/10426126.

[82] D. Tran *et al.*, "Functional genomics analysis reveals a MYC signature associated with a poor clinical prognosis in liposarcomas," *Am J Pathol,* vol. 185, no. 3, pp. 717-28, Mar 2015, doi: 10.1016/j.ajpath.2014.11.024.

[83] A. C. Tsiatis *et al.*, "Prognostic significance of c-Myc expression in soft tissue leiomyosarcoma," *Mod Pathol,* vol. 22, no. 11, pp. 1432-8, Nov 2009, doi: 10.1038/modpathol.2009.113.

[84] P. G. Nuciforo *et al.*, "Molecular and immunohistochemical analysis of HER2/neu oncogene in synovial sarcoma," *Hum Pathol,* vol. 34, no. 7, pp. 639-45, Jul 2003, doi: 10.1016/s0046-8177(03)00238-7.

[85] V. Barbashina *et al.*, "Oncoproteins and proliferation markers in synovial sarcomas: a clinicopathologic study of 19 cases," *J Cancer Res Clin Oncol,* vol. 128, no. 11, pp. 610-6, Nov 2002, doi: 10.1007/s00432-002-0389-3.

[86] A. Zoubek *et al.*, "Does expression of different EWS chimeric transcripts define clinically distinct risk groups of Ewing tumor patients?," *J Clin Oncol,* vol. 14, no. 4, pp. 1245-51, Apr 1996, doi: 10.1200/JCO.1996.14.4.1245.

[87] E. de Alava *et al.*, "EWS-FLI1 fusion transcript structure is an independent determinant of prognosis in Ewing's sarcoma," *J Clin Oncol,* vol. 16, no. 4, pp. 1248-55, Apr 1998, doi: 10.1200/JCO.1998.16.4.1248.

[88] A. Kawai, J. Woodruff, J. H. Healey, M. F. Brennan, C. R. Antonescu, and M. Ladanyi, "SYT-SSX gene fusion as a determinant of morphology and prognosis in synovial sarcoma," *N Engl J Med,* vol. 338, no. 3, pp. 153-60, Jan 15 1998, doi: 10.1056/NEJM199801153380303.

[89] H. Inagaki, T. Nagasaka, T. Otsuka, E. Sugiura, N. Nakashima, and T. Eimoto, "Association of SYT-SSX fusion types with proliferative activity and prognosis in synovial sarcoma," *Mod Pathol,* vol. 13, no. 5, pp. 482-8, May 2000, doi: 10.1038/modpathol.3880083.

[90] P. H. Sorensen *et al.*, "PAX3-FKHR and PAX7-FKHR gene fusions are prognostic indicators in alveolar rhabdomyosarcoma: a report from the children's oncology group," *J Clin Oncol,* vol. 20, no. 11, pp. 2672-9, Jun 1 2002, doi: 10.1200/JCO.2002.03.137.

[91] T. Saito *et al.*, "E-cadherin mutation and Snail overexpression as alternative mechanisms of E-cadherin inactivation in synovial sarcoma," *Oncogene,* vol. 23, no. 53, pp. 8629-38, Nov 11 2004, doi: 10.1038/sj.onc.1207960.

[92] T. Izumi *et al.*, "Prognostic significance of dysadherin expression in epithelioid sarcoma and its diagnostic utility in distinguishing epithelioid sarcoma from malignant rhabdoid tumor," *Mod Pathol,* vol. 19, no. 6, pp. 820-31, Jun 2006, doi: 10.1038/modpathol.3800599.

[93] T. Izumi *et al.*, "Dysadherin expression as a significant prognostic factor and as a determinant of histologic features in synovial sarcoma: special reference to its inverse relationship with E-cadherin expression," *Am J Surg Pathol,* vol. 31, no. 1, pp. 85-94, Jan 2007, doi: 10.1097/01.pas.0000213413.33558.85.

[94] O. Sato *et al.*, "Expression of epidermal growth factor receptor, ERBB2 and KIT in adult soft tissue sarcomas: a clinicopathologic study of 281 cases," *Cancer,* vol. 103, no. 9, pp. 1881-90, May 1 2005, doi: 10.1002/cncr.20986.

[95] Y. Oda, A. Sakamoto, T. Saito, N. Kinukawa, Y. Iwamoto, and M. Tsuneyoshi, "Expression of hepatocyte growth factor (HGF)/scatter factor and its receptor c-MET correlates with poor prognosis in synovial sarcoma," *Hum Pathol,* vol. 31, no. 2, pp. 185-92, Feb 2000, doi: 10.1016/s0046-8177(00)80218-x.

[96] D. P. Cahill *et al.*, "Mutations of mitotic checkpoint genes in human cancers," *Nature,* vol. 392, no. 6673, pp. 300-3, Mar 19 1998, doi: 10.1038/32688.

[97] G. J. Kops, D. R. Foltz, and D. W. Cleveland, "Lethality to human cancer cells through massive chromosome loss by inhibition of the mitotic checkpoint," *Proc Natl Acad Sci U S A,* vol. 101, no. 23, pp. 8699-704, Jun 8 2004, doi: 10.1073/pnas.0401142101.

[98] C. Kobayashi *et al.*, "Aberrant expression of CHFR in malignant peripheral nerve sheath tumors," *Mod Pathol,* vol. 19, no. 4, pp. 524-32, Apr 2006, doi: 10.1038/modpathol.3800548.

[99] Y. Oda *et al.*, "Altered expression of cell cycle regulators in myxofibrosarcoma, with special emphasis on their prognostic implications," *Hum Pathol,* vol. 34, no. 10, pp. 1035-42, Oct 2003, doi: 10.1053/s0046-8177(03)00404-0.

[100] H. Nakanishi, A. Myoui, T. Ochi, and K. Aozasa, "P-glycoprotein expression in soft-tissue sarcomas," *J Cancer Res Clin Oncol,* vol. 123, no. 6, pp. 352-6, 1997, doi: 10.1007/BF01438312.

[101] Y. Oda *et al.*, "Nuclear expression of Y-box-binding protein-1 correlates with P-glycoprotein and topoisomerase II alpha expression, and with poor prognosis in synovial sarcoma," *J Pathol,* vol. 199, no. 2, pp. 251-8, Feb 2003, doi: 10.1002/path.1282.

[102] C. Seidel *et al.*, "Alterations of cancer-related genes in soft tissue sarcomas: hypermethylation of RASSF1A is frequently detected in leiomyosarcoma and associated with poor prognosis in sarcoma," *Int J Cancer,* vol. 114, no. 3, pp. 442-7, Apr 10 2005, doi: 10.1002/ijc.20707.

[103] M. Cieslik and A. M. Chinnaiyan, "Cancer transcriptome profiling at the juncture of clinical translation," *Nat Rev Genet,* vol. 19, no. 2, pp. 93-109, Feb 2018, doi: 10.1038/nrg.2017.96.

[104] M. D'Errico *et al.*, "Genome-wide expression profile of sporadic gastric cancers with microsatellite instability," *Eur J Cancer,* vol. 45, no. 3, pp. 461-9, Feb 2009, doi: 10.1016/j.ejca.2008.10.032.

[105] C. M. Perou *et al.*, "Molecular portraits of human breast tumours," *Nature,* vol. 406, no. 6797, pp. 747-52, Aug 17 2000, doi: 10.1038/35021093.

[106] M. E. Garber *et al.*, "Diversity of gene expression in adenocarcinoma of the lung," *Proc Natl Acad Sci U S A,* vol. 98, no. 24, pp. 13784-9, Nov 20 2001, doi: 10.1073/pnas.241500798.

[107] X. Chen *et al.*, "Gene expression patterns in human liver cancers," *Mol Biol Cell,* vol. 13, no. 6, pp. 1929-39, Jun 2002, doi: 10.1091/mbc.02-02-0023.

[108] E. M. F. De Sousa *et al.*, "Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions," *Nat Med,* vol. 19, no. 5, pp. 614-8, May 2013, doi: 10.1038/nm.3174.

[109] R. Cristescu *et al.*, "Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes," *Nat Med,* vol. 21, no. 5, pp. 449-56, May 2015, doi: 10.1038/nm.3850.

[110] T. O. Nielsen *et al.*, "Molecular characterisation of soft tissue tumours: a gene expression study," *Lancet,* vol. 359, no. 9314, pp. 1301-7, Apr 13 2002, doi: 10.1016/S0140-6736(02)08270-3.

[111] S. Zhao, W. P. Fung-Leung, A. Bittner, K. Ngo, and X. Liu, "Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells," *PLoS One,* vol. 9, no. 1, p. e78644, 2014, doi: 10.1371/journal.pone.0078644.

[112] D. Delen, "Analysis of cancer data: a data mining approach," *Expert Systems,* vol. 26, no. 1, pp. 100-112, 2009, doi: https://doi.org/10.1111/j.1468-0394.2008.00480.x.

[113] D. B. Allison, X. Cui, G. P. Page, and M. Sabripour, "Microarray data analysis: from disarray to consolidation and consensus," *Nat Rev Genet,* vol. 7, no. 1, pp. 55-65, Jan 2006, doi: 10.1038/nrg1749.

[114] S. Paik *et al.*, "Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer," *J Clin Oncol,* vol. 24, no. 23, pp. 3726-34, Aug 10 2006, doi: 10.1200/JCO.2005.04.7985.

[115] S. Paik *et al.*, "A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer," *N Engl J Med,* vol. 351, no. 27, pp. 2817-26, Dec 30 2004, doi: 10.1056/NEJMoa041588.

[116] Y. Y. Syed, "Oncotype DX Breast Recurrence Score((R)): A Review of its Use in Early-Stage Breast Cancer," *Mol Diagn Ther,* vol. 24, no. 5, pp. 621-632, Oct 2020, doi: 10.1007/s40291-020-00482-7.

[117] L. Xie, L. Cai, F. Wang, L. Zhang, Q. Wang, and X. Guo, "Systematic Review of Prognostic Gene Signature in Gastric Cancer Patients," *Front Bioeng Biotechnol,* vol. 8, p. 805, 2020, doi: 10.3389/fbioe.2020.00805.

[118] J. Y. Cho *et al.*, "Gene expression signature-based prognostic risk score in gastric cancer," *Clin Cancer Res,* vol. 17, no. 7, pp. 1850-7, Apr 1 2011, doi: 10.1158/1078-0432.CCR-10-2180.

[119] A. E. Denton, E. W. Roberts, and D. T. Fearon, "Stromal Cells in the Tumor Microenvironment," *Adv Exp Med Biol,* vol. 1060, pp. 99-114, 2018, doi: 10.1007/978-3-319-78127-3_6.

[120] H. Wang, X. Wu, and Y. Chen, "Stromal-Immune Score-Based Gene Signature: A Prognosis Stratification Tool in Gastric Cancer," *Front Oncol,* vol. 9, p. 1212, 2019, doi: 10.3389/fonc.2019.01212.

[121] W. Dai *et al.*, "Identification of an EMT-Related Gene Signature for Predicting Overall Survival in Gastric Cancer," *Front Genet,* vol. 12, p. 661306, 2021, doi: 10.3389/fgene.2021.661306.

[122] F. Chibon *et al.*, "Validated prediction of clinical outcome in sarcomas and multiple types of cancer on the basis of a gene expression signature related to genome complexity," *Nat Med,* vol. 16, no. 7, pp. 781-7, Jul 2010, doi: 10.1038/nm.2174.

[123] E. H. Ren, Y. J. Deng, W. H. Yuan, Z. L. Wu, G. Z. Zhang, and Q. Q. Xie, "An immune-related gene signature for determining Ewing sarcoma prognosis based on machine learning," *J Cancer Res Clin Oncol,* vol. 147, no. 1, pp. 153-165, Jan 2021, doi: 10.1007/s00432-020-03396-3.

[124] Y. Shi *et al.*, "A risk signature-based on metastasis-associated genes to predict survival of patients with osteosarcoma," *J Cell Biochem,* vol. 121, no. 7, pp. 3479-3490, Jul 2020, doi: 10.1002/jcb.29622.

[125] W. Chung *et al.*, "Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer," *Nat Commun,* vol. 8, p. 15081, May 5 2017, doi: 10.1038/ncomms15081.

[126] K. Kim, S. O. Zakharkin, and D. B. Allison, "Expectations, validity, and reality in gene expression profiling," *J Clin Epidemiol,* vol. 63, no. 9, pp. 950-9, Sep 2010, doi: 10.1016/j.jclinepi.2010.02.018.

[127] V. P. Balachandran, M. Gonen, J. J. Smith, and R. P. DeMatteo, "Nomograms in oncology: more than meets the eye," *Lancet Oncol,* vol. 16, no. 4, pp. e173-80, Apr 2015, doi: 10.1016/S1470-2045(14)71116-7.

[128] A. Iasonos, D. Schrag, G. V. Raj, and K. S. Panageas, "How to build and interpret a nomogram for cancer prognosis," *J Clin Oncol,* vol. 26, no. 8, pp. 1364-70, Mar 10 2008, doi: 10.1200/JCO.2007.12.9791.

[129] C. Yu and Y. Zhang, "Development and validation of prognostic nomogram for young patients with gastric cancer," *Ann Transl Med,* vol. 7, no. 22, p. 641, Nov 2019, doi: 10.21037/atm.2019.10.77.

[130] D. S. Han *et al.*, "Nomogram predicting long-term survival after d2 gastrectomy for gastric cancer," *J Clin Oncol,* vol. 30, no. 31, pp. 3834-40, Nov 1 2012, doi: 10.1200/JCO.2012.41.8343.

[131] Q. Chen, L. Hu, and K. Chen, "Construction of a Nomogram Based on a Hypoxia-Related lncRNA Signature to Improve the Prediction of Gastric Cancer Prognosis," *Front Genet,* vol. 11, p. 570325, 2020, doi: 10.3389/fgene.2020.570325.

[132] Y. Bai *et al.*, "Development and Validation of a Prognostic Nomogram for Gastric Cancer Based on DNA Methylation-Driven Differentially Expressed Genes," *Int J Biol Sci,* vol. 16, no. 7, pp. 1153-1165, 2020, doi: 10.7150/ijbs.41587.

[133] Y. Liu *et al.*, "Development and validation of a hypoxia-immune-based microenvironment gene signature for risk stratification in gastric cancer," *J Transl Med,* vol. 18, no. 1, p. 201, May 14 2020, doi: 10.1186/s12967-020-02366-0.

[134] Y. Y. Song and Y. Lu, "Decision tree methods: applications for classification and prediction," *Shanghai Arch Psychiatry,* vol. 27, no. 2, pp. 130-5, Apr 25 2015, doi: 10.11919/j.issn.1002-0829.215044.

[135] F. J. Brims *et al.*, "A Novel Clinical Prediction Model for Prognosis in Malignant Pleural Mesothelioma Using Decision Tree Analysis," *J Thorac Oncol,* vol. 11, no. 4, pp. 573-82, Apr 2016, doi: 10.1016/j.jtho.2015.12.108.

[136] M. Takada *et al.*, "Prediction of axillary lymph node metastasis in primary breast cancer patients using a decision tree-based model," *BMC Med Inform Decis Mak,* vol. 12, p. 54, Jun 13 2012, doi: 10.1186/1472-6947-12-54.

[137] J. Sun *et al.*, "Development and validation of a hypoxia-related gene signature to predict overall survival in early-stage lung adenocarcinoma patients," *Ther Adv Med Oncol,* vol. 12, p. 1758835920937904, 2020, doi: 10.1177/1758835920937904.

[138] X. Bao, R. Shi, T. Zhao, and Y. Wang, "Immune landscape and a novel immunotherapy-related gene signature associated with clinical outcome in early-stage lung adenocarcinoma," *J Mol Med (Berl),* vol. 98, no. 6, pp. 805-818, Jun 2020, doi: 10.1007/s00109-020-01908-9.

[139] R. Shi *et al.*, "Establishment and Validation of an Individualized Cell Cycle Process-Related Gene Signature to Predict Cancer-Specific Survival in Patients with Bladder Cancer," *Cancers (Basel),* vol. 12, no. 5, May 2 2020, doi: 10.3390/cancers12051146.

[140] H. J. Leach, D. P. O'Connor, R. J. Simpson, H. S. Rifai, S. K. Mama, and R. E. Lee, "An exploratory decision tree analysis to predict cardiovascular disease risk in African American women," *Health Psychol,* vol. 35, no. 4, pp. 397-402, Apr 2016, doi: 10.1037/hea0000267.

[141] J. C. Chang *et al.*, "Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer," *Lancet,* vol. 362, no. 9381, pp. 362-9, Aug 2 2003, doi: 10.1016/S0140-6736(03)14023-8.

[142] J. Hannemann *et al.*, "Changes in gene expression associated with response to neoadjuvant chemotherapy in breast cancer," *J Clin Oncol,* vol. 23, no. 15, pp. 3331-42, May 20 2005, doi: 10.1200/JCO.2005.09.077.

[143] H. K. Kim *et al.*, "A gene expression signature of acquired chemoresistance to cisplatin and fluorouracil combination chemotherapy in gastric cancer patients," *PLoS One,* vol. 6, no. 2, p. e16694, Feb 18 2011, doi: 10.1371/journal.pone.0016694.

[144] J. Y. Douillard *et al.*, "Irinotecan combined with fluorouracil compared with fluorouracil alone as first-line treatment for metastatic colorectal cancer: a multicentre randomised trial," *Lancet,* vol. 355, no. 9209, pp. 1041-7, Mar 25 2000, doi: 10.1016/s0140-6736(00)02034-1.

[145] S. Giacchetti *et al.*, "Phase III multicenter randomized trial of oxaliplatin added to chronomodulated fluorouracil-leucovorin as first-line treatment of metastatic colorectal cancer," *J Clin Oncol,* vol. 18, no. 1, pp. 136-47, Jan 2000, doi: 10.1200/JCO.2000.18.1.136.

[146] K. R. Hess *et al.*, "Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer," *J Clin Oncol,* vol. 24, no. 26, pp. 4236-44, Sep 10 2006, doi: 10.1200/JCO.2006.05.6861.

[147] J. H. Cheong *et al.*, "Predictive test for chemotherapy response in resectable gastric cancer: a multi-cohort, retrospective analysis," *Lancet Oncol,* vol. 19, no. 5, pp. 629-638, May 2018, doi: 10.1016/S1470-2045(18)30108-6.

[148] A. Ribas *et al.*, "Association of response to programmed death receptor 1 (PD-1) blockade with pembrolizumab (MK-3475) with an interferon-inflammatory immune gene signature," *Journal of Clinical Oncology,* vol. 33, no. 15_suppl, pp. 3001-3001, 2015, doi: 10.1200/jco.2015.33.15_suppl.3001.

[149] L. Fehrenbacher *et al.*, "Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial," *Lancet,* vol. 387, no. 10030, pp. 1837-46, Apr 30 2016, doi: 10.1016/S0140-6736(16)00587-0.

[150] D. Zeng *et al.*, "Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures," *Cancer Immunol Res,* vol. 7, no. 5, pp. 737-750, May 2019, doi: 10.1158/2326-6066.CIR-18-0436.

[151] J. Zhang *et al.*, "Comprehensive characterization of the tumor microenvironment for assessing immunotherapy outcome in patients with head and neck squamous cell carcinoma," *Aging (Albany NY),* vol. 12, no. 22, pp. 22509-22526, Nov 18 2020, doi: 10.18632/aging.103460.

[152] R. Cao, L. Yuan, B. Ma, G. Wang, and Y. Tian, "Tumour microenvironment (TME) characterization identified prognosis and immunotherapy response in muscle-invasive bladder cancer (MIBC)," *Cancer Immunol Immunother,* vol. 70, no. 1, pp. 1-18, Jan 2021, doi: 10.1007/s00262-020-02649-x.

[153] Y. R. Miao *et al.*, "ImmuCellAI: A Unique Method for Comprehensive T-Cell Subsets Abundance Prediction and its Application in Cancer Immunotherapy," *Adv Sci (Weinh),* vol. 7, no. 7, p. 1902880, Apr 2020, doi: 10.1002/advs.201902880.

[154] F. Petitprez *et al.*, "B cells are associated with survival and immunotherapy response in sarcoma," *Nature,* vol. 577, no. 7791, pp. 556-560, Jan 2020, doi: 10.1038/s41586-019-1906-8.

[155] M. Albertsmeier *et al.*, "Cancer Testis Antigens and Immunotherapy: Expression of PRAME Is Associated with Prognosis in Soft Tissue Sarcoma," *Cancers (Basel),* vol. 12, no. 12, Dec 3 2020, doi: 10.3390/cancers12123612.

Acknowledgements

First of all, I would like to express my appreciation and thanks to Prof. Dr. med. Martin Angele for the tremendously productive environment, highly motivating scientific discussions and all the trust that he put in me.

I would like to express my sincere thanks to PD. Dr. med. Florian Bösch for his support, advice and trust in my work. I will not have my publications without his support. His enthusiasm for scientific research motivated me and from his excellent scientific knowledge I could learn a lot in the last years.

Special thanks to Prof. Dr. Alexandr Bazhin and Prof. Dr. med. Jens Neumman. They are always supportive of me throughout my time in Munich. I cannot make it without their supervision on my thesis and publications.

I would like to thank to my family for their support and faith in me during any time in my life and to my friends for motivating me and having enjoyable times out of the lab.

Finally, I gratefully acknowledge the financial support from China Scholarship Council.