



UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

EMBL



dkfz.

GERMAN
CANCER RESEARCH CENTER
IN THE HEIMHOLTZ ASSOCIATION

Research for a Life without Cancer



HEIDELBERG
UNIVERSITY
HOSPITAL

UMM

UNIVERSITÄTSMEDIZIN
MANNHEIM



AIHealth

INNOVATION CLUSTER

Funded Project Abstracts

| | |
|--|----------|
| Funded Project Abstracts..... | 1 |
| First Cohort..... | 4 |
| PostDoc Projects..... | 4 |
| AIH11: Inferring dynamics of clinically-relevant lipid markers of NASH and their master regulators from single-cell data..... | 4 |
| AIH12: Predicting tipping points towards sepsis in ICU patients using deep dynamical systems reconstruction..... | 4 |
| AIH20: Computational integration of cellular circuits and immune cell repertoires in health and disease..... | 5 |
| AIH21: Integration of Thermal Proteome Profiling and phospho-proteomics within network models to dissect disease mechanisms and drugs..... | 5 |
| AIH23: Detection and validation of somatic mosaicism of variants in amyotrophic lateral sclerosis..... | 6 |
| AIH28: Explainable Deep Neural Network-based Coronary Artery Disease Characterization via Photon-Counting..... | 6 |
| AIH31: PepAISim: Combining AI and molecular simulation for anticancer peptide and peptidomimetic design..... | 7 |
| AIH36: Model-based Artificial Intelligence in Surgical Data Science..... | 7 |
| Staff Scientist Projects..... | 8 |
| AIH01: Anonymous Synthesizer for Health Data (ASyH)..... | 8 |
| AIH03: Digital (Twin) Cousins Generator for Radiotherapy Research..... | 8 |
| AIH07: Machine-learning assisted detection of microsatellite instability..... | 9 |
| Clinician Scientist Projects..... | 10 |
| In All Beginnings Dwells a Magic Force: How Deep Learning Can Transform State of the Art Head and Neck Cancer Radiation Therapy..... | 10 |
| Machine Learning assisted early diagnosis of sepsis based on microcirculatory alterations in ICU patients..... | 10 |
| Creation of a standard dataset for laparoscopic Roux-en-Y gastric bypass procedures..... | 11 |
| Second Cohort..... | 12 |
| PostDoc Projects..... | 12 |
| AIH1: COmorbidity modeling IN spinal Cord Injury and DEpression (COINCIDE)..... | 12 |

| | |
|---|-----------|
| AIH4: Never split the difference - Machine Learning approaches to full-length transcript isoforms (ONT-SPLiT-seq) deconvolution..... | 12 |
| AIH6: Computational models of social interactions as a basis for adaptive gamified treatment approaches..... | 13 |
| AIH9: Estimating the health and economic burden induced by heatwaves in Germany using machine learning methods..... | 13 |
| AIH19: Combining Multiplexed Imaging and Computational Frameworks to Reveal Cellular Metabolic Interactions in the Human Tumor Microenvironment..... | 14 |
| AIH26: AI-guided design of functional RNA origami structures..... | 14 |
| Staff Scientist Projects..... | 15 |
| AIHII: Surgical AI platform – a core facility fostering translational research and clinical innovation for decision support in surgical oncology..... | 15 |
| AIHV: Generation of multiplex data and integration with electronic medical records for dimension reduction towards prediction of complications in critical illness..... | 15 |
| AIH14: Identification of plasma proteins causally linked to cancer development through integrated deep and machine learning..... | 16 |
| AIH21: Automated detection and molecular characterization of micronuclei in cancer cell lines and tumor tissues..... | 17 |
| AIH28: Gut microbiome profile as a diagnostic marker in eating disorders..... | 17 |
| Clinician Scientist Projects..... | 18 |
| AI-based risk prediction and diagnosis for heart failure and cardiomyopathy patients..... | 18 |
| Deep-learning based prediction of clinical outcome after endovascular therapy in acute ischemic stroke using topographic information of CT perfusion maps..... | 18 |
| Robust Deep Learning for CT-based Pedicle Screw Planning in Navigated Spinal Instrumentation..... | 19 |
| Xeno-Learning for Spectral Image Interpretation Surgery..... | 19 |
| Third Cohort..... | 20 |
| PostDoc Projects..... | 20 |
| AIH01: Identifying predictive dynamical signatures of anaesthesia effects using deep dynamical systems reconstruction..... | 20 |
| AIH04: Quantifying the mechanistic contribution of sex chromosomes and sex hormones to cell-type specific gene expression by genetic decoupling..... | 20 |
| AIH06: Machine-learning enhanced deep-tissue imaging for decoding neuron-glia interactions | 21 |
| AIH10: Integration of LightSeq data into single cell and spatial transcriptomic atlases of glioblastoma and neural development..... | 21 |
| AIH11: Bias-aware machine learning for automated sepsis diagnosis with hyperspectral imaging..... | 22 |
| Staff Scientist Projects..... | 23 |
| AIH I: Post-operative tissue fragment puzzle – How to improve patient care by solving a tissue-piece puzzle?..... | 23 |
| AIH III: AI-Powered Forensic Medical Examination: Improving Care for Children Affected by Violence..... | 23 |
| AIH V: Analysis of 3D Light-Sheet Microscopy Images of the Mouse Brain Using Deep Learning..... | 24 |
| AIH VI: AI for serological analysis of COVID-19 patients using multiplex microscopy assay.. | 24 |
| AIH VIII: Brain Tumor image and molecular data hub..... | 25 |
| AIH02: CatatonAI: deciphering molecular mechanisms in catatonia using multi-modal AI models..... | 25 |
| AIH03: Development of a Novel Digital Tool for Screening of Respiratory Illness in Children - Using an Interdisciplinary Approach..... | 26 |
| Clinician Scientist Projects..... | 27 |

| | |
|---|----|
| AIH01: Artificial intelligence (AI) based brain tumor classification for H&E-based and intraoperative diagnostics: Making advanced personalized diagnostics globally possible!..... | 27 |
| AIH02: Look Before You Leap: Radiotherapy Response Monitoring Using Deep Learning-Based Photoacoustic Image Analysis..... | 27 |
| AIH03: Joint AI-based analysis of motion patterns and white matter signatures for predicting treatment outcome in catatonia: A prospective bi-institutional study..... | 28 |
| AIH04: Using unsupervised clustering techniques to predict the kidney function based on a data fabric..... | 28 |
| AIH05: Using unsupervised clustering techniques to predict the kidney function based on a data fabric..... | 29 |
| AIH06: Data Science of Large Integrated Patient Data Sets for AI-Based Risk Prediction and Treatment Plans in General, Visceral And Transplant Surgery..... | 30 |
| AIH07: The Regulatory Role of miRNAs in Cardiovascular Diseases..... | 30 |
| AIH09: Development of a multimodal automated oxygen control device for resuscitation of preterm infants in the delivery room..... | 31 |
| AIH11: Longitudinal Hyperspectral Imaging (HSI) to Guide Intraoperative Hemodynamic Management..... | 31 |

First Cohort

PostDoc Projects

AIH11: Inferring dynamics of clinically-relevant lipid markers of NASH and their master regulators from single-cell data

Theodore Alexandrov, EMBL

Julio Saez-Rodriguez, Heidelberg University / UKHD

Abstract: Non-alcoholic steatohepatitis (NASH) is an inflammatory stage of the non-alcoholic fatty liver disease (NAFLD) and is a key factor of liver fibrosis, irreversible damage, and cancer. Novel molecular analyses of NASH are critically needed since no drugs are currently approved. Lipids were shown to be clinically-relevant markers discriminating healthy liver from NAFLD and NASH. However, interpretation of these markers is challenging in particular due to the limited understanding of their dynamics along the disease progression. Here, we will address this challenge by developing and applying new computational strategies for single-cell lipidomics. We will use these to investigate the in vitro model of NASH, recently obtained by the host team together with the Heikenwalder group at DKFZ. First, we will perform network analysis of lipids detected in the single-cell data and create a hypergraph linking co-detected lipids with the regulating enzymes. Next, we will infer dynamics of the clinically-known lipid markers of NASH in the single-cell data. Finally, we will find clusters of lipids co-regulated with the markers along the pseudo-time trajectory and infer dynamics of their master regulatory enzymes. Overall, this project will create a novel methodology with a potential to elucidate mechanisms of the reprogramming of lipid metabolism upon NASH and to yield novel targets for NASH therapies.

AIH12: Predicting tipping points towards sepsis in ICU patients using deep dynamical systems reconstruction

Daniel Durstewitz, ZI

Verena Schneider-Lindner, UMM

Holger Lindner, UMM

Abstract: Sepsis is the leading cause of death in the intensive care unit (ICU). Empirically, sepsis requires rapid and specific therapy, but its bedside diagnosis still largely relies on patient assessment by clinical experts. Building on a large ground truth database of ICU patients that we have been collecting since 2016, we aim to develop novel AI approaches based on interpretable recurrent neural networks designed for nonlinear dynamical systems (DS) reconstruction. These methods yield subject-specific DS models which can be mathematically analyzed using DS theory tools, and can be simulated under different scenarios. Such models will hence not only enable onsite running forward predictions of individual patient trajectories, but will also allow for advanced identification of future tipping

points, crucial transitions (bifurcations) in the underlying disease dynamics toward sepsis. Moreover, through their interpretability from a DS perspective, we will gain mechanistic insight into the pathogenic processes that could be harvested for better tailored clinical interventions. It is therefore expected that the development and validation of this approach will have major clinical implications and could be integrated into clinical decision support systems. The planned studies crucially rely on the complementary medical-statistical and AI & DS expertise of the collaborating groups.

AIH20: Computational integration of cellular circuits and immune cell repertoires in health and disease

Julio Saez-Rodriguez, UKHD

Hedda Waredemann, DKFZ

Lucas Schirmer, Heidelberg University / UMM

Abstract: Understanding the functioning and regulation of the immune system is key to treating a range of conditions from autoimmune diseases to cancer, as well as improving vaccinations. Recent advances in single-cell technologies provide molecular data with unprecedented detail, including with spatial or time resolution, that is information on where individual cells are placed in a tissue or when they develop over time, respectively. This project aims to link two currently decoupled areas in computational biology, analysis of immune repertoire and characterization of molecular circuits, and apply them to single-cell resolved data to study the effect of immune repertoires on intra-and inter-cellular signaling. Using as illustrative cases own spatially-resolved data sets obtained in the context of chronic inflammation in tissue of patients and time-resolved data following vaccine responses, we expect to improve our understanding of the deregulation of the immune system, paving the way for improved treatments of a variety of conditions.

AIH21: Integration of Thermal Proteome Profiling and phospho-proteomics within network models to dissect disease mechanisms and drugs

Mikhail Savitski, EMBL

Julio Saez-Rodriguez, Heidelberg University / UKHD

Abstract: Understanding the mode of action of drugs as well as the molecular underpinnings of disease states is essential for developing new therapeutic strategies. Powerful proteomics technologies such as phosphoproteomics as well as the recently developed thermal proteome profiling (TPP) hold the potential for deciphering and understanding cellular processes that are modulated by drug treatment and altered in disease. Currently, there are no adequate tools for comprehensively extracting the relevant information from TPP experiments alone or in combination with phosphoproteomics. We propose to leverage the expertise present in our labs and to develop powerful new tools that will be capable of mining these datasets thus

elucidating drug mode of action, as well as assessing cellular changes specific to cell lines that represent different disease states. We are convinced that this computational framework will become instrumental in multiple efforts within the Alliance. All tools developed in the context of this project will be released as free, easy-to-use open software and become available to the scientific community.

AIH23: Detection and validation of somatic mosaicism of variants in amyotrophic lateral sclerosis

Lena Voith von Voithenberg, DKFZ

Benedikt von Brors, DKFZ

Rossana Parlato & Jochen Weishaupt, UMM

Abstract: Amyotrophic lateral sclerosis (ALS) is a fatal degenerative disease of the nervous system. Only few percent of ALS cases are familial with a known underlying germline mutation, whereas 90% of the cases occur spontaneously. In this project, we will investigate the hypothesis that somatic mutations in few cells (somatic mosaicism) play a role in the development of sporadic ALS. To this end, we will analyse somatic variants by targeted sequencing in a panel of known risk genes in post-mortem tissue of the motor cortex and spinal cord of ALS patients. To determine whether somatic mosaicism of variants exists, we will differentiate between low frequency variants and sequencing artefacts by using error correction based on unique molecular identifiers. We will further establish and implement combinations of AI-supported variant calling and filtering strategies with differing feature selection strategies and will validate candidate genes by using amplicon sequencing and mutation-specific in situ hybridization approaches.

AIH28: Explainable Deep Neural Network-based Coronary Artery Disease Characterization via Photon-Counting

Sandy Engelhardt, UKHD

Stefan Schönberg, UMM

Abstract: Photon counting Computed Tomography (CT) is a major technological advancement in the field of radiology. It allows for significantly better resolved images by reduced radiation at the same time. While this is an already existing development, especially for imaging small scale structures like coronary plaque, tissue characterization might become feasible with other contrast agents at high resolution. This will have a disruptive impact in the field of cardiovascular diseases, where 3D image interpretation is mostly hampered by high MRI slice thicknesses (e.g. short axis acquisitions of 8 mm voxel length). As part of the PC3 consortium (www.pc3.eu) UMM has one of the first photon counting CT machines in Germany installed.

In this proposal, based on our previous work and available infrastructure, we form a novel collaboration and aim to address two issues: We aim to develop novel deep learning based methods for plaque characterization and assessment of severity of stenosis. This combined

With the concept of explainability and quantitative and qualitative quality control, as supervised learning techniques require highly standardized annotations and rigorous testing. Furthermore, the novel tissue resolving capabilities of gadolinium-enhanced CT will be investigated and correlated with MRI to assess fibrosis and myocardial scar formation.

AIH31: PepAISim: Combining AI and molecular simulation for anticancer peptide and peptidomimetic design

Rebecca Wade, Heidelberg University
Elke Burgermeister, UMM

Abstract: The goal of this project is to leverage significant recent advances in AI methods for the design and mechanistic characterization of peptides and peptidomimetics for therapeutic purposes with a focus on cancer. As a case study, we will focus on the potential for treating gastrointestinal cancers with peptides that mimic the effects of the tumor suppressor “myotubularin-related protein 7” (MTMR7). Computational approaches, combining machine learning with molecular simulation, will be used to investigate the protein-protein interactions of MTMR7, design peptides, and predict their molecular mechanisms. State-of-the-art methods employing deep learning and invertible neural networks that are being developed in Heidelberg and elsewhere will be applied. Computations will guide experiments to functionally characterize the interaction candidates and the experiments will inform further computations in a tight, iterative process. Peptides will be assessed for anti-tumor efficacy in patient-derived tumor organoid models and their molecular mechanisms of action will be explored in human gastrointestinal cancer cells using in vitro protein interaction assays. We expect this project to advance AI methods for drug discovery and their practical application to the discovery of therapeutic peptides, and to lead to the discovery of specific molecular mechanisms and therapeutics against gastrointestinal cancers.

AIH36: Model-based Artificial Intelligence in Surgical Data Science

Lena Maier-Hein, DKFZ
Beat Müller, UKHD

Abstract: Death within 30 days after surgery is the third-leading cause of death worldwide, with research suggesting that a large proportion of these deaths are due to surgical error. The newly established domain of Surgical Data Science (SDS) aims to improve the quality and outcome of interventional healthcare through the capturing, organizing, analyzing and modeling of data. However, clinical translation of data science methods proves difficult. A large international consortium of experts recently attributed the lack of clinical success stories to the lack of large annotated databases. Data sparsity can thus be regarded as the main roadblock in the field of SDS. Previous approaches have addressed this roadblock with a diverse set of methods including crowdsourcing, self-supervised learning, active learning

and synthetic data generation. In this project, we investigate an entirely complementary approach based on integrating existing medical knowledge in neural network-based analysis. Specifically, we propose the integration of prior knowledge encoded in ontologies in a Graph Neural Network (GNN) based-approach to surgical decision support. The method will be validated based on a data set from the TIGER study, an international effort to include 5000 patients with esophageal cancer, and two in-house data sets on esophageal cancer and intrahepatic Cholangioma.

Staff Scientist Projects

AIH01: Anonymous Synthesizer for Health Data (ASyH)

Christoph Dieterich, UKHD

Martin Lablans, DKFZ

Motivation: Our ASyH proposal aims to remove barriers between rapid developments in the machine learning community and the inert world of hospital information management concepts and systems. On one hand, any AI-based machine learning approach requires adequate data. On the other hand, all data originating in the health sector are rightfully protected from uncontrolled use and access.

Approach: For this specific situation in the health community, we propose to generate suitable data sets through AI-based anonymization as well as data synthetization methods. We take up the challenge of generating non-identifiable data according to GDPR regulations. To this end, we use available data in the university hospitals of Mannheim and Heidelberg to build on our own and use existing AI models for data de-identification and data synthesis. These models preserve the underlying statistical properties of the original data. Another important aspect in ML is data augmentation as sufficient data is oftentimes scarce.

Expected outcome: Our proposal has a clear service-oriented character as it caters AI-ready data from local hospitals to the machine learning community. Herein, we use our own and existing software solutions and local infrastructure (including MII DICs). We will also evaluate a very relevant question: To what extent are data generated at both partner sites actually comparable?

AIH03: Digital (Twin) Cousins Generator for Radiotherapy Research

Oliver Jäkel, DKFZ

Jens Fleckenstein, UMM

Kristina Gisk, DKFZ

Abstract: With AI on the rise in radiotherapy, our research community is facing an increasing challenge to source enough patient scans for scrutinizing deep learning models for segmentation, registration, image synthesis, and dose prediction along adaptive treatment courses. Specifically, pilot-phase projects suffer from delays due to ethics vote application, volunteer recruitment, and missing secure transfer channels from clinical systems to research teams. Sharing these medical data is even more challenging. Open access datasets seldom provide all necessary scan types to investigate adaptive radiotherapy research questions. Especially in radiotherapy, scans from multiple modalities acquired over a month-long treatment course are needed to assess accumulated dosimetric effects. The aim of the proposed project is to establish a hybrid - AI and rule-based - computational infrastructure able to generate artificial anatomies mimicking radiotherapy patients. GAN-like random anatomy generators shall be combined with rule-based biomechanical motion models and imaging simulators to create representative artificial patient cohorts for development of technical AI prototypes in the radiotherapy community. With this, also research teams around the world without access to clinical databases could progress scientific insight by publishing proposed developments alongside the utilized artificial data or by benchmarking their proposed solutions to earlier approaches on the same artificial data cohorts.

AIH07: Machine-learning assisted detection of microsatellite instability

Matthias Kloor, UKHD

Fruzsina Molnár-Gábor, Heidelberg University

Abstract: Microsatellite instability (MSI) is one of the two major types of genomic instability driving human cancers. Cancers with MSI develop because of a functional defect in their DNA mismatch repair system and occur in the colorectum (15% of colorectal cancers), the endometrium (25-30%) and a broad variety of other malignancies. The clinical impact of MSI is wide-ranging: MSI tumor patients (1) show limited benefit from standard chemotherapy, (2) have a better prognosis, and (3) are generally highly responsive to immunotherapy. Therefore, immune checkpoint blockade has been approved by the FDA for all MSI cancers irrespective of tumor origin, representing the first basket admission entirely based on a molecular marker. Despite the high and rapidly increasing significance of MSI testing, the high costs, technical hurdles and complexity of the analysis hampers implementation of MSI reflex testing in routine diagnostic procedures. Recent independent studies have demonstrated that prediction of the MSI phenotype by machine learning-assisted classification of histopathology images is highly promising. The present project intends (1) to establish the data life cycle for a multi-center initiative compliant with data protection law to optimize machine learning-assisted prediction of MSI and (2) to prepare roll-out of AI classification algorithms to clinical application in diagnostics.

Clinician Scientist Projects

In All Beginnings Dwells a Magic Force: How Deep Learning Can Transform State of the Art Head and Neck Cancer Radiation Therapy

Sebastian Adeberg, UKHD
Klaus Maier-Hein, DKFZ

Abstract: Due to technical advances, the outcome has improved significantly in head and neck cancer radiation therapy. But local tumor recurrence occurs often and the quality of life of patients is impaired by treatment toxicities. Moreover, treatment planning remains complex and time-consuming. Thus, the two main research aims of the project are (I) to reduce treatment planning time and target volume variability using autosegmentation to create consistent, high-quality target volumes and (II) to determine tumor control probabilities to enhance outcome prediction and clinical decision making. The work program will be based on training and cross-validation data sets derived from retrospective cohorts that will be further evaluated as part of prospective clinical trials. All data will be made available in the radiation oncology and data science communities according to common open science principles. By reducing the time required for treatment planning, the efficiency and flexibility of current clinical workflows can be transformed, thereby saving costs and improving consistency and treatment quality. By modeling of tumor control and normal tissue complication probabilities, patient selection, treatment plan optimization and personalized radiation treatments will be enhanced. Thereby, the current developments can improve the oncological outcome and reduce treatment toxicities for the benefit of the patients.

Machine Learning assisted early diagnosis of sepsis based on microcirculatory alterations in ICU patients

Markus A. Weigand, UKHD
Lena Maier-Hein, DKFZ
Annette Kopp-Schneider, DKFZ

Abstract: Sepsis is a life-threatening systemic response to infection leading to organ failure, with an associated 8% increase in mortality per hour of delay of therapy. Microcirculatory alterations are a fundamental component of septic organ failure. Currently, no objective microcirculatory monitoring is established in the clinical routine of the intensive care unit (ICU). Hyperspectral imaging (HSI) is an innovative technique which allows for fast and non-invasive microcirculatory measurements, without risk or relevant burden to the patient. In a first pilot study, HSI of skin microcirculation has shown potential for sepsis detection. However, to date, no data from larger cohorts of patients with microcirculatory abnormalities measured with HSI is available. As the worldwide first ICU, we will establish microcirculatory HSI measurements as routine monitoring. Furthermore, we aim to implement a registry of all

ICU patients with subsequent development and validation of a machine learning algorithm for automated sepsis diagnosis based on the acquired data. We expect machine learning assisted evaluation of HSI to open up new opportunities in early diagnosis and initiation of sepsis therapy, thus improving survival.

Creation of a standard dataset for laparoscopic Roux-en-Y gastric bypass procedures

Christoph Reissfelder, UMM

Jürgen Hesser, UMM

Lena Maier-Hein, DKFZ

Abstract: The application of data science in surgery has been far less successful than in other fields of science. An international consortium of experts has identified missing datasets as the major hurdle in enabling Surgical Data Science (SDS)¹. To create high-quality datasets, data needs to be collected according to strict acquisition protocols and annotated by trained personnel following strict annotation protocols. This requires significant effort and the access to expert knowledge represents a major bottleneck. By creating, curating and publishing a dataset for laparoscopic Roux-en-Y gastric bypass, an operation that requires performing intestinal anastomosis, we aim to facilitate the use of SDS methods in training and during operations. The collected data will be two-fold, consisting of (1) data deriving from training sessions and (2) real-life procedures in operating rooms (OR). Care will be taken to systematically capture and annotate perioperative data (Surgical team, skill level, patient data from the Electronic Health Record (EHR)), video clips of surgeons using standardized protocols that will be published together with the dataset. Modern tools will be used for annotation², and algorithm specific metadata (Implementation, Version, Uncertainty, Training data used, etc.) will be collected and included in the dataset.

Second Cohort

PostDoc Projects

AIH1: COmorbidity modeling IN spinal Cord Injury and DEpression (COINCIDE)

Emanuel Schwarz, ZI
Norbert Weidner, UKHD

Abstract: Depression is a highly frequent and severely disabling condition affecting approximately 40% to 60% of patients with traumatic spinal cord injury (SCI). It has serious consequences for the short- and long-term rehabilitation of individuals with SCI, and is associated with an increased suicide risk. However, the biology of depression in SCI is poorly understood and no biological tools exist that could aid in identifying at-risk patients. COINCIDE, an interdisciplinary project focusing the development and application of artificial intelligence methods to characterize the mechanisms underlying depression in SCI, aims to address this and provide the basis for the development of novel diagnostic and predictive tools. The project will develop and apply multi-task machine learning approaches to characterize the link between neurological SCI deficits and depression-relevant brain function. The project builds on already available data from SCI patients, as well as large-scale brain-functional data acquired in healthy controls, as well as patients with major depressive disorders. COINCIDE will then determine genetic and epigenetic associations in the same individuals, to characterize the biology of depression in SCI, and provide the basis for the development of novel diagnostic and predictive tools.

AIH4: Never split the difference - Machine Learning approaches to full-length transcript isoforms (ONT-SPLiT-seq) deconvolution

Christoph Dieterich, Heidelberg University
Lars Steinmetz, EMBL

Abstract: Mutations in the cardiac splice factor RBM20 cause severe dilated cardiomyopathy (DCM) however, due to technological limitations, its regulatory network is only partially understood. SPLiT-seq is one of the few single-cell sequencing methods for the analysis of cardiomyocytes. In several split-pool rounds, fixed cells are randomly distributed into wells and transcripts are labelled with well-specific barcodes. The Steinmetz Lab (EMBL/Stanford) and Dieterich Lab (U. Heidelberg) combine their expertise on Bioinformatics, long-read Nanopore sequencing, RNA splicing and systems cardiology to rethink SPLiT-seq completely new. While conventional SPLiT-seq uses short Illumina sequencing, we will propel SPLiT-seq into the era of long-read, single-molecule Nanopore sequencing without compromising on accuracy and speed. We have recently established a combined protocol for SPLiT-seq and

long-read sequencing (ONT-SPLiT-seq), which we propose to employ for the generation of a single cell isoform atlas of the murine heart. Next, we will identify cell populations that are prone to mis-splicing in mice with patient-relevant mutations in RBM20. Finally, we will integrate for the first time CRISPR-mediated perturbations with single-cell long read analysis to associate gene function to splicing in a high-throughput fashion. Combined, this approach will provide crucial insights into the RBM20-mediated splice network and its perturbation in DCM.

AIH6: Computational models of social interactions as a basis for adaptive gamified treatment approaches

Georgia Koppe, ZI
Christoph Korn, UKHD

Abstract: Social interactions form a key aspect of our everyday life. Through cooperative exchanges with our fellow human beings, we obtain behaviorally relevant information, as well as emotional support and protection. Disturbed or altered interaction behavior is a core feature in multiple psychiatric disorders, and is perceived as extremely burdensome and stressful, both by the affected individuals as well as by those in their immediate surroundings. Here, we propose to adopt modern Machine Learning (ML) algorithms from the reinforcement learning (RL) domain to model and emulate (alterations in) social interaction behavior in healthy and diseased individuals. Highly predictive models will then be applied to develop model-based adaptive gamified treatment approaches to modify individuals' behavior, as well as to gain mechanistic insights into the pathogenic processes that underlie maladaptive social interactions and their neural substrates. By encouraging positive social learning experiences, the proposed approach is hypothesized to promote learning of more advantageous interaction strategies, enhance cooperation, and strengthen core social competences. We expect that the development and validation of this ML approach will have strong clinical implications and could be integrated into treatment approaches for psychiatric disorders.

AIH9: Estimating the health and economic burden induced by heatwaves in Germany using machine learning methods

Manuela De Allegri, UKHD
Joacim Rocklöv, Heidelberg University
Alina Hermann, UKHD

Abstract: Climate change represents a major threat to human health, for instance due to heatwaves and other extreme weather events. In 2003 alone, about 70.000 excess deaths were attributed to heat within few weeks in Europe. The Lancet Countdown on health and climate change estimates the economic burden of heat induced excess mortality to correspond to 0.28% of the gross world product, with the greatest economic burden occurring in Europe. However, these costs are likely underestimated since very limited evidence is

available as to what costs heat imposes on the health system, due to increased health service utilisation, and on society at large, due to reduced productivity and premature mortality. Our study addresses this knowledge gap by combining economic evaluation methods with machine learning methods, to estimate the societal cost of heatwaves in Germany for 2011-2021. To achieve our objective, we examine how machine learning approaches perform compared to the state-of-the-art time series regression approaches to estimate the impact and predict costs by linking nationally representative insurance data with historic temperature data. This represents a unique attempt to reconcile different disciplinary approaches to produce more valid and credible results than those derived by standard approaches and to promote methodological advances.

AIH19: Combining Multiplexed Imaging and Computational Frameworks to Reveal Cellular Metabolic Interactions in the Human Tumor Microenvironment

Felix Hartmann, DKFZ
Julio Saez-Rodriguez, Heidelberg University

Abstract: Significant transition points in human cancer span tumor initiation, expansion, metastasis, and therapeutic resistance. Importantly, these transitions involve complex interactions between cells within the tumor microenvironment (TME). Novel spatial-omics technologies now provide an opportunity to interrogate this complexity at unprecedented resolution. In this project, we will combine a novel proteomic imaging platform (multiplexed ion beam imaging: MIBI) with machine learning frameworks to extract clinically significant cellular interactions from human melanoma tissue. MIBI visualizes the expression and spatial distributions of up to 40 proteins and thus enables deep analysis of the TME. Extracting knowledge from multiplexed spatial data requires the development of scalable computational methods that can leverage the availability of the spatial context. We have recently developed such approaches, with an emphasis on leveraging prior biological knowledge to extract mechanistic insights. Bringing together this novel cutting-edge imaging technology with AI-supported data analysis frameworks will advance our understanding of cellular interactions and their functional consequences and may thus lead to novel biomarkers and potential therapeutic targets for human cancer.

AIH26: AI-guided design of functional RNA origami structures

Kerstin Göpfrich, Max Planck Institute for Medical Research
Frauke Gräter, Heidelberg University

Abstract: Bottom-up synthetic biology and DNA/RNA nanotechnology are two distinct frontier fields. What unites them is their rational engineering mindset, their common understanding that precise function needs precise components – with the aim to either build a cell or to repurpose nucleic acids for nanoscale construction. Here, we will explore how AI-guided

evolutionary approaches can lead to a step change in the functional complexity of DNA/RNA-based nanomachines and synthetic cells alike because they allow us to explore a large design space. In particular we will probe RNA-lipid interactions with AI-accelerated multi-scale simulations. This alone will provide insights on the potential role of RNA-lipid interactions at the origins of life. Moreover the simulations yield design instructions for functional RNA origami structures, in particular (i) RNA nanopores and (ii) vesicle division inducing RNA structures. These structures will be implemented experimentally and tested in lipid vesicles as synthetic cellular compartments. We thereby envision an interdisciplinary contribution to the field of bottom-up synthetic biology as well as applicable RNA origami-based tools for nanopore sensing and as genetically encoded biophysical probes in cell biology.

Staff Scientist Projects

AIHII: Surgical AI platform – a core facility fostering translational research and clinical innovation for decision support in surgical oncology.

Martin Dugas, UKHD
Martin Wagner, UKHD
Lena Maier-Hein, DKFZ

Abstract: Our AI innovation is to provide a platform for clinical translation of data driven decision support in surgical oncology. We aim to link data from the medical data integration center at UKHD with novel surgical data science algorithms developed at the DKFZ to be used by surgeon scientists in the labs and operating rooms of the surgical department. In particular, the requested staff scientist will maintain and further develop the Kaapana platform that has been developed at DKFZ within the NCT-DSDSO-project to meet the crucial need for a platform within the surgical department of UKHD. Firstly, the platform will be used to extract multimodal surgical data from their primary sources and integrate it into the semantic data model of the MEDIC-environment of UKHD in order to obtain semantically enriched structured data in a standardized format (FHIR, OpenEHR, ODM). Secondly, algorithms developed at DKFZ will be integrated into the real-time environment of the operating rooms within the surgical department. Finally, in order to facilitate surgical AI innovation for a larger community we will establish a core facility within the surgical department to support both, other clinician AI scientists at UKHD and basic scientist inside and outside DKFZ and UKHD.

AIHV: Generation of multiplex data and integration with electronic medical records for dimension reduction towards prediction of complications in critical illness

Holger Lindner, UMM
Lars Feuerbach, DKFZ

Abstract: The immune response to severe injury (trauma, infection, cancer...) and its treatment critically determines the risk of acute and difficult-to-predict clinical complications. Among these, sepsis and organ dysfunction predominate. New prognostic and diagnostic classifiers are urgently sought to improve patient outcomes. Electronic medical records (EMRs) in the intensive care unit (ICU) capture clinical courses at high resolution. Their integration with large-scale longitudinal immunophenotyping has potential to significantly advance classifier discovery by AI. To this end, the staff scientist will direct medium-throughput multiplex analyses of blood markers of immunity in 1100 available blood samples collected from 100 ICU patients and leverage expertise in processing EMR data at UMM and in application of AI for nonlinear dimensionality-reduction at DKFZ. Multiplex data will be integrated with EMR data from same patients and time points and analysed by uniform manifold approximation and projection (UMAP). Ground truth labels on sepsis and organ dysfunction will be projected onto resulting UMAP embeddings. We seek clusters associated with the onset of clinical complications and afferent trajectories of patient time for analyses of the clinical features/blood markers that essentially define them. We expect these features/markers to inform future patient stratification strategies and to represent potential classifiers for further investigations.

AIH14: Identification of plasma proteins causally linked to cancer development through integrated deep and machine learning

Jeroen Krijgsveld, DKFZ
Justo Lorenzo Bermejo, UKHD

Abstract: Establishing causal relationships between plasma protein levels and cancer development is essential for understanding disease aetiology and improving current cancer prevention.

Problematically, differences in plasma protein levels between cases and healthy controls may anticipate disease (protein→cancer) or arise as a consequence of the tumour (cancer→protein). Single nucleotide polymorphisms (SNPs) associated with plasma protein levels can be used to infer directionality, distinguishing proteins that precede disease and thus

inform risk prediction and precision prevention. We aim to apply deep and machine learning to a unique collection of prospective plasma samples from 104 gallbladder cancer case-control

pairs with extensive epidemiological and SNP data. After generating mass-spectrometry (MS) proteomics data for this cohort, we will use deep neural networks for differential plasma

proteome profiling based on MS features. We will then apply advanced machine learning

techniques (robust LASSO) to predict plasma protein levels based on individual SNP data, finally leading to novel “protein→cancer” associations. Collectively, the sequential application

of two AI tools (deep and machine learning) to newly generated proteomics data for a unique collection of prospective plasma samples, and to large existing genetic-proteomic cancer data, will provide a blueprint towards risk prediction for other disease entities using different types of molecular data.

AIH21: Automated detection and molecular characterization of micronuclei in cancer cell lines and tumor tissues.

Denis Schapiro, UKHD
Fred Hamprecht, Heidelberg University

Abstract: Chromosomal instability (CIN) is a hallmark of cancer and manifests as structural or numerical alterations in the chromosomes. CIN is associated with widespread therapeutic resistance, immune evasion, and metastasis. Despite its prominent role, no druggable molecular targets are currently known. To enable a systematic assessment of CIN, we will leverage genetic constructs of members of the kinesin family to generate various levels of chromosomally unstable cell lines. The most common way to assess CIN is by micronucleus scoring, which is usually performed manually by counting micronuclei or abnormalities under a microscope. Here we propose to (i) create a computational approach to automatically count micronuclei and associated abnormalities in immunofluorescence images. Next, (ii) we will perform spatial omics analysis on those cell lines to create a molecular profile of CIN. Finally, (iii) we will apply the here developed computational methods to human tissue samples (connected to clinical information) to uncover potential druggable molecular targets for CIN. The spatial omics technologies and single cell analysis pipelines created by AG Schapiro combined with the image analysis and machine learning capabilities by AG Hamprecht provide a unique setup for this interdisciplinary project.

AIH28: Gut microbiome profile as a diagnostic marker in eating disorders.

Hans-Christoph Friederich, UKHD
Georg Zeller, EMBL

Abstract: The gut microbiota is essential in the regulation of appetite and body weight and has been found to be altered in eating disorders. Previous research has identified gut microbiota dysbiosis as a potential pathognomonic aspect of Anorexia Nervosa, Bulimia Nervosa and Binge Eating Disorder. Changes in microbiota composition and production of bioactive metabolites may contribute to the development and/or maintenance of eating disorders. In this project, we plan to longitudinally assess gut microbiota composition and function in patients with eating disorders during in-house treatment to gain a better understanding of microbiome changes associated with progression or remission of eating disorders. In line with similar investigations in obesity, where variations in the gut microbiome

allow classification of individuals according to their risk of further weight gain, we aim to establish an AI-based diagnostic procedure for the classification of patients into different profiles of disease progression.

Clinician Scientist Projects

AI-based risk prediction and diagnosis for heart failure and cardiomyopathy patients

Benjamin Meder, UKHD
Klaus Maier-Hein, DKFZ

Abstract: Cardiovascular diseases are one of the leading causes of death in industrialized nations. In Germany, heart failure (HF) is also the most common reason for hospitalization in adults. This cardiovascular pandemic not only burdens the patients and their families, but also strains the resources of the health system. In view of the increasingly aging population, this burden will continue to increase in the future. We propose an AI-based risk prediction for heart failure and cardiomyopathy patients to address this problem with a scalable approach. AI can serve cardiologists through a clinical decision support system (CDSS) and even enhance diagnosis and risk prediction by evaluating complex patterns. In this project, we aim to combine clinical data from multimodal sources, routinely acquired in the current standard patient workup in the clinics, to create a comprehensive AI model for everyday clinical practice in cardiology.

Deep-learning based prediction of clinical outcome after endovascular therapy in acute ischemic stroke using topographic information of CT perfusion maps

Ulf Neuberger, UKHD
David Bonekamp, DKFZ

Abstract: Stroke is a leading cause of mortality and disability worldwide, with a global lifetime risk of approximately 25%. A major factor determining where patient might benefit from interventional recanalization is the ischemic infarction that may already present at the timepoint of treatment selection. Perfusion imaging is well suited to delineate permanently infarcted areas from tissue-at-risk, which can possibly be saved with successful treatment. However, this approach only accounts for the volume of the tissue-at-risk, but not for the function of the damaged or endangered tissue. In our project, we aim to overcome this limitation by applying a deep-learning based approach, that will not only consider the volume of ischemic tissue and tissue-at-risk, but also the topographic distribution, the severity of ischemia and the specific symptoms of patients who have been successfully treated with an

endovascular technique in our institution. The findings of this project will be used to improve selection of patients who would benefit from endovascular therapy for acute ischemic stroke.

Robust Deep Learning for CT-based Pedicle Screw Planning in Navigated Spinal Instrumentation

Moritz Scherer, UKHD
Klaus Maier-Hein, DKFZ

Abstract: Navigation and robotic systems have been increasingly applied in spine surgery but dedicated screw planning is a time-consuming prerequisite to tap the full potential of these guidance techniques. We recently described a deep-learning (DL)-based approach to automatic lumbar screw planning which produced robust results that were non-inferior to manual planning by spine experts holding great potential to improve time-efficacy in surgical workflows when converted into fully equipped surgical planning tool. Our goal is to extend the applicability of the DL-approach to thoracic spine segments requiring additional training of the algorithm and refinement of the model as thoracic vertebrae differ critically in size and pedicular anatomy from previously leveraged lumbar vertebrae. An iterative learning approach is sought enable a constant improvement of the algorithm incorporating expert-based corrections of screw suggestions prior to implantation during surgery. Lastly, maximization of screw dimensions and consideration of local bone quality for adjustment of screw trajectories according to maximized screw fastening strength can contribute to a reduction of implant failures due to poor bone quality and osteoporosis in a constantly aging patient population requiring spine surgery nowadays.

Xeno-Learning for Spectral Image Interpretation Surgery

Felix Nickel, UKHD
Lena Maier-Hein, DKFZ

Abstract: Death within 30 days after surgery has been found to be the third-leading cause of death worldwide¹. One of the major challenges faced by the surgeons is the visual discrimination and evaluation of tissues. To overcome the limitations of visual perception, hyperspectral imaging (HSI) might present a solution. While conventional medical cameras are limited by “imitating” the human eye; hyperspectral cameras remove this arbitrary restriction. Instead, they capture multiple bands of light that decode relevant information on tissue type and perfusion. The bottleneck related to converting the potential of this novel imaging technique into patient benefit is related to the lack of large annotated human data sets. Consequently, this project aims to address this issue based on a new concept that we refer to as *Xeno-Learning*. The core idea is to address the shortage of data by transferring knowledge from one species (specifically porcine) to another (here: human). The core methodological data science challenge is to develop a data representation that enables the generalization not only to new individuals but also to other species. To this end, we can

leverage a huge data set consisting of HSI data comprising several thousand images of animals and humans.

Third Cohort

PostDoc Projects

AIH01: Identifying predictive dynamical signatures of anaesthesia effects using deep dynamical systems reconstruction

Daniel Durstewitz, CIMH
Simon Wiegert, UMM

Abstract: Invasive surgical procedures often require general anaesthesia, but many common anaesthetics may cause unwanted and prolonged side effects such as post-operative delirium and memory loss. To contain these and other potential issues during surgery, online monitoring of anaesthesia depth is vital. It is furthermore important to understand which pharmacological agents, under which conditions, achieve effective anaesthesia, while at the same time avoiding harmful side effects. The present project aims at identifying dynamical signatures of brain activity that allow to predict the clinical effectiveness of anaesthetics and their dosage as well as their later impact on memory processes and consciousness. We will use deep recurrent neural networks (RNNs), with architectures and training procedures optimized for nonlinear dynamical systems reconstruction, to extract such signatures from large-scale Ca^{2+} imaging and extracellular electrophysiological measurements from rodent hippocampus under the impact of different anaesthetics, administered at various doses. The dynamical profiles identified, e.g. complexity and temporal structure of limit cycles, will then be used as predictors for effectiveness and side effects. Trained RNNs will furthermore be used to gain mechanistic insight into the connections between network dynamics, anaesthesia depth, and effects on memory consolidation.

AIH04: Quantifying the mechanistic contribution of sex chromosomes and sex hormones to cell-type specific gene expression by genetic decoupling

Moritz Gerstung, DKFZ
Edith Heard, EMBL
Duncan Odom, DKFZ

Abstract: In all mammals, including human, there are transcriptional differences between the sexes that have profound impact on organismal homeostasis and disease. During the covid pandemic, for instance, men were far more likely to die following coronavirus infection, reflecting a profound immunological difference with women; however, the mechanistic origins of these sex differences remain poorly understood. This project will generate large-scale quantitative data from millions of single-cell transcriptomes by exploiting a novel genetic mouse model that decouples sex chromosomes (XX v XY) from an individual's genetically engineered sex (male v female). With this model we are generating data of sufficient scale to quantitatively resolve the relative contribution of sex hormones and sex chromosomes to cell-specific transcription. To measure the relative contributions of sex hormones and chromosomes across the entire transcriptome using data from millions of cells, we will adopt start-of-the-art machine learning and pattern recognition algorithms. This interdisciplinary approach will additionally afford insights into what role, if any, X-chromosome escapee genes play in cell-type-specific regulatory programs.

AIH06: Machine-learning enhanced deep-tissue imaging for decoding neuron-glia interactions

Amit Agarwal, HU

Robert Prevedel, EMBL

Anna Kreshuk, EMBL

Abstract: Our brain consists of an equal proportion of broadly classified groups of cells called neurons and glia. Although glial cells constitute 50% of the brain, the mechanisms by which glia connect and communicate with the neurons in health and disease remain unknown. To understand neuron-glia interactions, we require novel non-invasive optophysiological technologies, which enable investigation of the structure and function of glial cells in intact brains. We have recently developed intravital deep-tissue multiphoton microscopy with adaptive optics which can deliver volumetric images at sufficiently high speed and resolution. The bottleneck in the analysis of complex cellular communication in vivo now lies with the image analysis part of the pipeline which needs to extract the signals from large, noisy image volumes, at a high throughput rate. The aim of our proposal is to bring together AI-based computer vision techniques and methods of computational neuroscience to reconstruct the cell morphology and calcium signals in an accurate and reliable manner. Furthermore, we will make the resulting tools accessible to a wider neuroscience community. AI-driven computational analysis methods will play a key role in our multidisciplinary approach to study physiology and pathophysiology of the neurons and glia in the intact brain

AIH10: Integration of LightSeq data into single cell and spatial transcriptomic atlases of glioblastoma and neural development

Julio Saez-Rodriguez, UKHD

Sinem Saka, EMBL

Varun, Venkataramani, DKFZ

Abstract: Glioblastoma is an incurable brain cancer characterised by its invasiveness and notorious therapeutic resistance. The underlying cellular and molecular heterogeneity of glioblastoma has been increasingly characterised with single-cell RNA-sequencing and spatial transcriptomics. In contrast, the functional relevance of various glioblastoma cell subpopulations is not yet understood. We recently developed a workflow that goes from functional live imaging to transcriptome-wide spatial characterisation to allow a simultaneous characterisation of cellular behaviour in glioblastoma. The scope of this proposed project is (i) to integrate the spatially-defined transcriptomics data acquired by applying Light-Seq on co-culture models, animal models and patient samples into existing single cell and spatial transcriptomic atlases of glioblastoma using various deep learning strategies, (ii) to jointly analyze the data to characterize the key multi-cellular processes that define this devastating tumour, and (iii) to uncover neurodevelopmental trajectories of glioblastoma by integrating these datasets into atlases of neural development. Taken together, machine learning approaches to interrogate these complex multimodal datasets will significantly increase our ability to identify novel therapeutic targets.

AIH11: Bias-aware machine learning for automated sepsis diagnosis with hyperspectral imaging

Lena Maier-Hein, DKFZ

Markus A. Weigand, UKHD

Abstract: Sepsis is one of the leading causes of death and critical illness, accounting for approximately 19.7 % of all global deaths in 2017. Early sepsis detection is a key factor in patient recovery because the mortality rate increases with every hour the antimicrobial intervention is delayed. Despite decades of clinical research, sepsis detection is limited to few (usually unreliable) biomarkers. Although a recent study found characteristic spectral signatures for septic and non-septic patients, automated machine learning-based diagnosis of sepsis from HSI data was hampered due to confounding factors being present in the data. Similarly, an increasing number of recent papers have revealed severe flaws in study design

that lead to confounded algorithms. To address the issue of confounding biases in sepsis diagnosis based on HSI, we propose to develop a method capable of transforming confounded HSI measurements into a so-called confounder-invariant space, where the influence of confounding variables on sepsis diagnosis is minimized. Based on our expertise in the analysis of confounders in HSI data, we will build an extensive database comprising HSI and digital patient data of septic and nonseptic patients. This will serve as the basis for our approach to confounder-invariant representation of HSI data for sepsis diagnosis.

Staff Scientist Projects

AIH I: Post-operative tissue fragment puzzle – How to improve patient care by solving a tissue-piece puzzle?

Cleo-Aron Weis, UKHD

Claudia Scherl, UKHD

Karl Rohr, Heidelberg University

Abstract: Thinking of puzzles, you probably quickly think of a straightforward child's play? For computers, puzzling is well-known as a complex task, especially when it comes to irregularly shaped human tissue fragments with artifacts and sometimes missing pieces in between, as in our case. In this project, surgical specimens from head and neck surgery will be reconstructed based on histological sections at the end of a complex work-up process. Determining exact anatomical situations is difficult and error-prone. Therefore, the aim is to determine the tumor size and location more correctly, in particular, to provide more reliable safety margins because patients with positive resection margins have a 2.5-fold increased risk of dying [1]. First, the specimen or puzzle fragment borders are determined using unsupervised CNN-based segmentation. Second, the borders are compared by sequence comparison (in analogy to Stanco et al. [2]), and a graph-based approach determines the best puzzle piece combination. Third, the tissue cross-sections reconstructed in this way are reconstructed in 3D. Finally, a 3D histological data set of a surgical specimen with additional resection specimens will be created, allowing to navigate and measure as in a radiological image data set.

AIH III: AI-Powered Forensic Medical Examination: Improving Care for Children Affected by Violence

Kathrin Yen, UKHD

Jürgen Hesser, UMM

Susanne Krömker, Heidelberg University

Abstract: The ARMED (Augmented Reality assisted, Forensic Medical Evidence collection and Documentation) project has been launched by the Institute for Forensic Medicine at the Heidelberg University Hospital to guide forensic examinations performed in associated hospitals. Using telemedicine and augmented reality technology, the examining physicians wear data glasses that transmit the findings to an expert in Heidelberg. The expert makes an expert assessment and provides the doctor on site with instructions on how to carry out the examination and document evidence. The project aims to create comprehensive and expert care for children affected by violence. The goal is to expand the service to other clinics and population groups and develop an AI-driven data analysis procedure. The images transmitted during the examination will be compared to the existing database, making suggestions for the possible origin of injuries. The structured reporting of findings can be used for comprehensive analysis, data mining, and research projects. The AI-driven process offers benefits such as cost-saving, more time for human interaction, and easy translation into other languages.

AIH V: Analysis of 3D Light-Sheet Microscopy Images of the Mouse Brain Using Deep Learning

Sevin Turcan, UKHD

Karl Rohr, Heidelberg University

Abstract: Brain tumors interact with the host brain, which allows these cells to grow and become resistant to therapies. Whole-brain 3D imaging in mouse models provides an opportunity to visualize and quantify tumor-host interactions. Light-sheet microscopy is a valuable tool since it allows rapid acquisition of high-resolution 3D images of whole tumor-bearing mouse brains. To understand brain tumor development, registration of the 3D image data of a subject with an atlas and statistical analysis of the distribution of cells are important. This project focuses on images with pathologies, which are more challenging compared to healthy brains used in previous work. Deep learning methods and software will be developed for registration of 3D light-sheet fluorescence microscopy images of the mouse brain with an atlas as well as for statistical analysis of the cell distribution within relevant brain regions.

AIH VI: AI for serological analysis of COVID-19 patients using multiplex microscopy assay

Vibor Laketa, UKHD

Anna Kreshuk, EMBL

Constantin Pape, University of Göttingen

Abstract: The emergence of the novel pathogenic coronavirus SARS-CoV-2 and its rapid pandemic spread had dramatic consequences on human society across the globe. Continuous evolution of new viral variants affecting infectivity, disease severity and immune evasion has challenged society's efforts to contain the virus. Mutations in the viral spike protein are of special interest, as all currently licensed vaccines against SARS-CoV-2 are based on the immune response to the spike. Early in the pandemic, we have established a microscopy-based assay which allowed for studies of population prevalence of the original strain of the virus. We now extend this work to simultaneously measure the levels of patient serum antibodies against different spike mutants and nucleocapsids. The assay was also applied for screening human therapeutic monoclonal antibodies. In the next phase we plan to detect auto-recognizing antibodies in patient sera as major determinants of Long COVID. The aim of this project is to develop an image analysis pipeline based on state-of-the-art AI-based methods for segmentation and classification of complex cellular staining patterns. The pipeline needs to be implemented as a flexible solution to be employed in other clinically.

AIH VIII: Brain Tumor image and molecular data hub

Felix Sahm, UKHD

Moritz Gerstung, DKFZ

Melanie Janning, UMM

Miriam Ratliff, UMM

Abstract: Histology and molecular data are indispensable in oncology. Today, transformative novel approaches to merge these two data layers exist. First, molecular context to tissue sections and morphology can be obtained by spatial sequencing technologies. Second, histology data can be leveraged to train machine learning algorithms that predict molecular characteristics based on H&E images alone. For the field of brain tumor research, neuropathology is at the center of these endeavors. Thus, a variety of data are generated in joint projects with our collaboration partners, including digitalized images, sequencing, array and clinical data. Maintenance and expansion of this data repository is of utmost importance for the collaborative projects in image-based marker prediction (with the Gerstung lab at DKFZ) or deciphering intra-tumoral niches of brain metastases (with the Janning and Ratliff group at UMM). A staff researcher can here leverage their current knowledge in data science of assessing high throughput data and grow into the field of integrative data science across multiple platforms.

AIH02: CatatonAI: deciphering molecular mechanisms in catatonia using multi-modal AI models

Carl Herrmann, UKHD

Dusan Hirjak, ZI

Abstract: Psychiatric conditions represent a diagnostic challenge, as many clinical manifestations are shared between otherwise distinct diseases, with different molecular alterations. Catatonia, a severe neuropsychiatric disorder with affective, motor and behavioural symptoms, is frequently associated with mental disorders, suggesting a shared underlying mechanism. Long classified as a subtype of schizophrenia, it is now considered a separate diagnostic entity. Neuroimaging, genetic and clinical data suggest that specific pathways, for example the GABAergic, dopaminergic and glutamatergic pathways, are affected. Using integrative deep-learning approaches based on explainable AI models, we intend to integrate the available data (genetic, transcriptomic and imaging) from specific catatonia studies with large scale datasets available for psychiatric conditions such as schizophrenia using transfer learning approaches. We identify altered mechanisms in catatonia patients and identify changes in pathway activities comparing longitudinal data from acutely ill and remitted patients. Finally, we will use these catatonia-specific disease signatures to retrospectively investigate schizophrenia cohorts to predict the probability of catatonic symptoms.

AIH03: Development of a Novel Digital Tool for Screening of Respiratory Illness in Children - Using an Interdisciplinary Approach

Claudia Denking, UKHD

Joachim Rocklöv, Heidelberg University

Jens Fleckenstein, UMM

Abstract: About 50-65% of children suffering from tuberculosis (TB) across all age groups are never detected. This contributes to the 250,000 deaths from TB in children annually, making the disease one of the top ten killers of children below the age of five. To improve TB detection in children and reduce TB-related morbidity and mortality, novel diagnostic tests are urgently needed. This project aims to improve case finding at all levels of care in resource-limited, high TB burden settings through the development and validation of an innovative, digital TB screening tool that provides individualized TB disease risk predictions. Starting with large geographically and demographically representative data sets from clinical trials and prevalence surveys of TB in children, we will develop predictive models using innovative supervised machine learning ensemble-based methods. We aim to increase explainability using game theoretical methods. We will implement the predictive algorithms in a web-tool for use by health personnel to determine risk of TB and optimize usability and trust through mixed methods research. With an efficient and effective tool in hand, we would subsequently seek funding for an international consortium to evaluate the design-locked digital tool in clinical case finding studies to inform WHO recommendations.

Clinician Scientist Projects

AIH01: Artificial intelligence (AI) based brain tumor classification for H&E-based and intraoperative diagnostics: Making advanced personalized diagnostics globally possible!

Felix Sahm, UKHD

Moritz Gerstung, DKFZ

Miriam Ratliff, UMM

Abstract: Neuropathology work-up of neurosurgical specimen is often costly and time-consuming due to the variety of molecular markers tested and methods involved. As a result, treatment is delayed and patients at centers without access to these technologies suffer from lower standards of care. Recent advances in molecular classification based on simple H&E images promise to overcome this limitation. To this end, large cohorts of intraoperative preparations and H&E images annotated with molecular data are essential. Likewise, interdisciplinary collaboration between bioinformaticians and neuropathologists is indispensable. In this project, collaboration of the clinician scientist fellow with bioinformatics staff at Dept. of Neuropathology HD, the experts in image-based marker prediction of the AI in Oncology group at DKFZ, and the neurosurgery colleagues at UMM Mannheim will devise a novel approach to comprehensive diagnostic with minimal resources in an instant manner.

AIH02: Look Before You Leap: Radiotherapy Response Monitoring Using Deep Learning-Based Photoacoustic Image Analysis

Thomas Held, UKHD

Lena Maier-Hein, DKFZ

Abstract: Multispectral optoacoustic tomography (MSOT) is a non-invasive, low-cost and dose-free real-time imaging method for analyzing functional tissue parameters such as oxygenation. First devices have recently been approved for clinical use thus opening new pathways for clinical applications. In radiotherapy, for example, monitoring of tumor progress over the course of several irradiations can potentially be improved and simplified with such a modality. In particular, the monitoring of tumor hypoxia, which has been repeatedly associated with increased radioresistance and unfavorable prognosis, promises to become feasible. To date, however, analyzing the high-dimensional, multispectral photoacoustic imaging data remains challenging. First approaches utilizing deep learning techniques show promise but still face challenges related to the typically limited, application-specific data sets. In particular, confounding factors directly related to the patient, the acquisition or the intervention are often not considered, which renders robust and reliable application of deep learning models difficult as causal relationships in the data cannot be correctly identified. In this project, the first

comprehensive photoacoustic data set of head & neck radiotherapy patients will be acquired and used to design deep learning models aiming to automatically analyze the radiotherapeutic treatment progress. Besides the pure photoacoustic imaging data, meta data concerning the patient, acquisition and intervention as well as high quality

semantic annotations of the target structures will be provided. For the first time, this will allow for development of deep learning solutions capable of providing confounder-invariant estimations of the treatment response throughout radiotherapy. We expect that this can ultimately lead to dynamic and personalized treatment plan optimization and improvement of the oncological outcome.

AIH03: Joint AI-based analysis of motion patterns and white matter signatures for predicting treatment outcome in catatonia: A prospective bi-institutional study

Dusan Hirjak, ZI

Peter Neher, DKFZ

Abstract: Catatonia is a serious psychiatric disorder that can be characterized by different motor, affective and cognitive-behavioural symptoms. Although previous studies have contributed to a better understanding of catatonia, these studies used heterogenous clinical rating scales, which are prone to subjective error measurements. Innovative 3D optical markerless motion capture system (MOCAP) enables clinical/behavioural data assessment in real-time, objectively and accurately, thereby overcoming challenges of clinical rating scales such as inter-observer variability and insensitivity. Drawing from the longitudinal whiteCAT study conducted by both PIs, this project follows three main objectives: First, we will use MOCAP to obtain multiparametric information regarding the spatiotemporal biomechanical characteristics of catatonia symptoms in a sample of 60 catatonia patients according ICD-11. Second, we will unravel the relationship between behavioural and neurobiological mechanisms underlying different catatonia symptoms. Third, we will use artificial intelligence (AI) models combining biomechanical and white matter parameters based on diffusion MRI (dMRI) to predict catatonia symptoms severity at baseline and after 12 weeks of follow-up. This interdisciplinary longitudinal project will test the feasibility of AI models consisting of clinical/behavioural and neuroimaging data in order to provide health practitioners with accurate and objective information regarding their catatonia patients.

AIH04: Using unsupervised clustering techniques to predict the kidney function based on a data fabric

Cleo-Aron Weis , UKHD

Florian Kälble, UKHD

Fabian Siegel, UMM

Abstract: Patients with chronic kidney failure usually have a long clinical story with heterogenous underlying diseases and multiple diagnostic procedures. Despite this information, through the course of the disease, clinically the prognosis is often difficult to predict and both prognostic and predictive markers are lacking. Against this background, the proposed project aims to retrospectively identify potential biomarkers to accurately predict the trajectory of a patient's kidney function over time. Therefore, a data fabric is created that combines a wide range of patient data, including laboratory results, imaging studies, medical history, and demographic factors. By means of unsupervised clustering methods, the diverse sub-datasets are labelled and mapped to a second knowledge graph containing the clinical course of the kidney function. Based on this second knowledge graph, graph-based machine-learning methods are applied to identify potential biomarkers. The main goal of this study is to identify patients at high risk of developing chronic kidney disease or progressing to end-stage renal disease. By identifying these patients early, healthcare providers could implement targeted interventions and treatment plans to slow or prevent the progression of kidney disease.

AIH05: Using unsupervised clustering techniques to predict the kidney function based on a data fabric

Christoph Reissfelder, UMM
Lena Maier-Hein, DKFZ

Abstract: During surgical training, objective feedback can significantly help the residents learn technical skills through performance assessment. Several scoring systems have been developed to evaluate the trainees' performance based on proficiency criteria, for example, the Objective Structured Assessment of Technical Skills (OSATS). Traditionally, using OSATS requires expert observers to evaluate trainees in their performance and is thus time-consuming and expensive. Moreover, manual assessment is subjective by nature. In this project, we aim to develop an artificial intelligence (AI)-based system for automated skill assessment, which can be used automatically for evaluating laparoscopic hand-sewn gastrojejunostomy based on OSATS scores. For this purpose, a standard dataset for laparoscopic Roux-en-Y gastric bypass procedures we created in previous work needs to be annotated following strict annotation protocols, which are to be developed in the project.

Some challenges exist in annotating surgical videos in surgical data science (SDS)¹. The inter-rater variability is high since most assessments can be influenced by the surgical experience and expectations of the annotators. Access to expert knowledge represents a major bottleneck: skilled surgeons are expensive, both financially and in terms of opportunity cost (time spent watching videos against time spent treating patients).

AIH06: Data Science of Large Integrated Patient Data Sets for AI-Based Risk Prediction and Treatment Plans in General, Visceral And Transplant Surgery

Christoph Michalski, UKHD

Martin Dugas, UKHd

Alba Diaz-Muñoz, EMBL

Abstract: The proposed project aims to improve data management and science in the University of Heidelberg Surgical Hospital and the affiliated European Pancreas Centre, including a standardised and digitalised system for streamlined patient care and high-quality research. The project's development involves stages such as developing a digital platform using the OpenEDC infrastructure, homogenising and standardising data acquisition protocols, improving patient pathways, and developing data interfaces between the hospital and science. This interdisciplinary project involves the collaboration of clinical and research teams in surgery, radiology, anaesthesiology, molecular biology and biophysics, IT experts, data scientists, and specialists of artificial intelligence (AI). The project also aims to establish an AI- powered data analysis system that integrates translational research data for personalised therapy, including patient questionnaire data and basic science data. The project aims to generate a unique dataset to enable new answers to essential clinical questions that require big data analysis. The proposed Clinician Scientist Fellow has a broad expertise in general, visceral, and transplantation surgery and the project leverages previous work on developing online data platforms and data management and integration in clinical and research settings.

AIH07: The Regulatory Role of miRNAs in Cardiovascular Diseases

Elham Kayvanpour, UKHD

Kai Jan, DKFZ

Oliver Stegle, DKFZ

Abstract: Cardiovascular disease (CVD) is a leading cause of morbidity and mortality. There is a pressing need for a more profound understanding of the disease and the identification of novel prognostic and diagnostic biomarkers for the prevention and treatment of CVD. With the emergence of precision medicine, microRNAs have demonstrated to be promising biomarkers. We have collected multicentric data from a large cohort of over 2000 patients and controls, and obtained microarray data for approximately 2000 miRNAs. The proposed project aims to evaluate the diagnostic potential of miRNA signatures using machine learning approaches for a range of cardiovascular diseases, including coronary artery disease, heart

failure with reduced ejection fraction, dilated cardiomyopathy, and acute coronary syndrome with a special focus on validating miRNAs that have been previously identified in the literature and investigating their impact on patient survival. Our findings may provide valuable information for the development of novel precision risk stratification tools and new therapies in CVD.

AIH09: Development of a multimodal automated oxygen control device for resuscitation of preterm infants in the delivery room

Christian Gille, UKHD

Jan Stallkamp, Heidelberg University

Abstract: In very preterm infants respiratory support including supplemental oxygen (O₂) during delivery room (DR) resuscitation is challenging. The aim is to maintain optimal oxygenation through adjustment of supplemental O₂ concentration guided by haemoglobin O₂ saturation (SpO₂) and heart rate (HR) values. The optimal O₂ titration strategy during the transition phase after birth is still lacking (1, 2). A recently DR practice audit has identified O₂ titration strategies as one of the most important areas for further improvement in neonatal resuscitation (3). Current O₂ titration strategies result in infants spending up to 60% of time outside of the SpO₂ target range (4, 5). Beyond the transition period automatic oxygen control (AOC) have been proven to improve oxygenation (6). However, during challenging resuscitations frequent manual titration steps are impractical. Therefore, we aim to set the basis to develop AOC devices for DR resuscitation by (a) providing a training dataset of optimized manual O₂ control in very preterm infants, (b) based on this to develop an AOC algorithm using AI, c) perform in- vitro simulation testing of the derived AOC algorithm on a validation dataset (d) clinical testing using the algorithm to further optimize manual adjustments in the delivery room (open loop manual adjustments).

AIH11: Longitudinal Hyperspectral Imaging (HSI) to Guide Intraoperative Hemodynamic Management

Jan Larmann, UKHD

Lena Maier-Hein, DKFZ

Rosa Klotz, UKHD

Abstract: Annually, approximately 4.2 million patients die following surgery. The major underlying complications are related to impaired microcirculatory tissue perfusion, leading to organ failure, and increased mortality. The lack of reliable methods to measure microcirculatory perfusion makes hemodynamic management challenging, resulting in severe morbidity and mortality. Hyperspectral imaging (HSI) is an innovative technique that offers

non-invasive, direct measurement of tissue oxygenation, perfusion, and tissue water content, potentially serving as a method for live examination of microcirculatory perfusion. The goal of this multidisciplinary AI project is to pioneer the use of deep learning-based analysis of HSI to optimize intraoperative hemodynamic management, maintain adequate organ perfusion during surgery, prevent complications, and improve outcomes. The project aims to establish a database of annotated HSI skin and organ images linked to clinical data, identify factors associated with insufficient microcirculatory perfusion, and develop a deep learning-based algorithm to predict organ failure and complications based on HSI. Machine learning will be used to identify factors associated with insufficient microcirculation and develop a protocol to guide hemodynamic management, which can be tested in subsequent large-scale clinical studies. The newly established collaboration between DKFZ and UKHD will establish a technology to guide intraoperative microcirculation management and optimize outcomes.
