Subtitles and transcriptions

Subtitles and transcriptions are available for selected materials for purpose of helping users understand the contents of the educational sessions.

Uncertain words have been indicated with ?? before and after the part. Parts that could not be understood at all have been indicated as [Audio Not Clear].

Every effort has been made to faithfully reproduce the audio of the sessions as recorded. However, no responsibility is accepted for mistakes or omissions. ESO does not endorse any opinions expressed in the presentations.

Colon cancer - using digital tools to search for old and new biomarkers

Dr Kather: Good morning, everybody, and welcome to my talk. My name is Jakob Kather. I'm a medical doctor in Aachen, in Germany. I focus on GI cancer and at the same time, I have a technical background in medical image analysis, and I lead a research group about digital image biomarkers in oncology, mostly in GI oncology. So, I'm very happy that in the next 20 minutes, I can talk to you about colon cancer, using digital tools to search for old and new biomarkers. Let's get started. These are my disclosures. And, okay, digital tools for new biomarkers. I will focus on artificial intelligence-based tools today. And whenever we talk about artificial intelligence in medicine, we actually talk about image analysis. So, why is this so? Well, the vast majority of all AI products that are available in medicine today are somehow related to image analysis. And digital images have just a lot of information that we can extract with artificial intelligence, especially with deep convolutional neural networks, and then we can make this information clinically usable. And interestingly analysing digital image data, was really hard for a long time. So before artificial intelligence tools became stable and widely used, people had a hard time to programme computers to detect objects in images and classify images automatically, especially in medicine. For example, detecting a melanoma, just based on a cell phone photograph of your skin, telling apart a melanoma and a benign skin lesion was a really hard task. It was actually not sufficiently solvable by computer-based image analysis tools before deep learning. And then all of these things changed in 2012, deep learning for deep convolutional neural networks, technology that powers this field, well, for the first time, outperformed, all other computational methods. And then a few years later, this trend also spilled over to medicine. And then in 2017, we saw this paper, which you can see here, where, for the first time in a large-scale study, deep learning was used to classify images of skin lesions. And in this particular dataset, the computer performed better than a panel of human experts. And this paper really started a big trend in this field. But of course, this technology is not just applicable to images of skin lesions, but also to any other type of medical image data. And here I show you for a comparison, a few data types that we deal with in medical oncology. So, for example, dermoscopy images, colonoscopy images, radiology images and histopathology images. And in this case, these images are shown at the same scale. So, one pixel in one image is the same size as one pixel in the other image. And what you can see here, is that basically, histopathology images are much, much larger than the other imaging types, when it comes to the number of pixels. So, the number of pixels in histology image is more than in a single slide, is more than 10 times the size of a whole chest CT series, as you can see here. So, that's why in our work, where we focus on extracting biomarkers from image data, we work a lot with histology images. However, the fact that these images are really large, is also a problem because we cannot just take them and put them into deep neural networks as they are, because they are so large, and our computer hardware cannot handle this. And that's why we have to take these large images, we have to cut them into smaller patches or tiles, and then we use these tiles and pre-process them, maybe normalise them so to remove some differences in staining.

And then we use these to train the deep neural networks. So how does our standard workflow look like? Well, in any machine learning project, what you have is basically a two-step approach. And this is you first train a classifier, and then you apply a classifier. So, what I'm showing you here, and so these approaches accounts for the vast majority of machine learning analysis and supervised approaches. Supervised means you have a training data set where you know which classes the instances belong to. For example, you have 10,000 training images, and 5,000 of them are a tumour and 5,000 of them are non-tumour images. And so, these are the labels, so we have the images and the labels, and we use these two pieces of information to train a deep neural network in the computer. And this what we call a supervised prediction problem, supervised approach. And this is really what most people do with AI in medicine. There are also other approaches, for example, unsupervised learning, where you don't have to have these labels and the network will discover the underlying structure in a data set by itself. And then there's also other approaches, for example, reinforcement learning, which you can use to teach a computer how to play chess, for example. But these are not so widely used at this point in medicine. Okay, so in our approach, in the supervised approach, we train the network with ground truth data that we have available, and then we can apply it to other images, where we don't know the label. We don't know which class they belong to, and the network will give us a prediction and also a probability of each image, belonging to each class. And for example, we can, this was applied for the photographs of the skin lesions, but we can also apply this to any other imaging data. And actually, this two-step approach, training and application, is not really hard. So, in this paper, in 2015, they trained actually pigeons to classify images of pathology and radiology of breast cancer into different classes, tumour, and non-tumour, mostly. They showed they could train those pigeons to do that. So, this simple classification is not so hard, but what is really hard is to make use of this, to apply this in a clinically relevant context. So last year what we did, was to somehow create the systematic overview of AI technology applications, in the context of cancer histopathology, because just a lot of work is happening in the histopathology application, but of course there's also is somehow true for other imaging types. So, we identified four types of studies. One type of study that just checked if tumour was present on the tissue slide or not. And so, for example, this is clinically relevant when you have biopsies of the prostate, where you suspect a tumour could be, and then you want to detect tiny, tiny pieces of tumour, a tiny glance in these biopsies. And this is where deep learning can really can really help. And many studies have shown that this works pretty well. Well, there's another type of study that you can see here in the right and top corner, that prediction of molecular alterations from tumour tissue. So, this is something I am especially interested in because we know that in precision oncology, it's very important to know the mutations in a given tumour to prescribe the tailored treatment for this patient. But we cannot sequence every patient, every tumour in every patient in the whole world, currently, because it's kind of expensive and it needs also dedicated infrastructure. So, our idea was to use these H&E images, to predict the probability of a mutation, a specific mutation being present, in this particular tumour. And this is something that we as a research group focus on, but also of course, other groups are working on this, and this works surprisingly well. So, for a number of tumour types, you can predict probabilities of mutations, directly from the H&E image. Other types of studies focus on prediction of overall survival from H&E images, and most importantly, or most relevantly, is prediction of treatment response directly from images. So, for example, there are some studies that have tried to predict the response to immunotherapy from histology images. And this worked and these proved very preliminary studies worked not too badly. So now, let me focus a little bit more on the prediction of mutations from the images. So, this is something that's still experimental, we don't have a medical product, medical device, diagnostic device approved right now, but a number of commercial, in addition to the academic groups working on this, a number of commercial companies are working on this and trying to bring a product to the market. So how would this look like? Well, currently our workflow in the clinic, if we take a piece of a tissue from a patient with a suspected tumour, and then part of it goes for histopathology and part of it goes for genetic testing. And then in the end, these pieces of information are put together and treatment recommendations are made. So, the idea is that maybe this workflow could be simplified by predicting the probabilities of mutations directly from H&E images and thereby having to run the genetic test only on a

subset of cases. So, the biological idea behind this is that, in any organism and in any tumour, there's a set of specific genetic makeup, and this genotype determines the phenotype. And the idea is that by looking at the phenotype you can basically guess the genotype or give a probability for a specific genotype. And the fact that this works to some degree in some tumours, has been shown by a number of publications. First of all, in 2018, by this by Coudray et al, who showed in lung cancer that you can predict mutational status from histology. Our own group has shown this in 2019 in colorectal cancer. And the first application here was to predict microsatellite instability, which is a clinically approved biomarker for cancer immunotherapy in metastatic colon cancer. So, and we could predict this MSI, microsatellite instability here, directly from H&E slides. And this was somehow an easy task, because it's already known that MSI leaves a specific morphological pattern in these H&E images. However, until then, these patterns were just not strong enough or clear enough to be used for diagnostic and really diagnostic purposes. And in a head-to-head comparison by a collaborating group of ours, it was also shown subsequently that the computer-based prediction of MSI status is much better than the human-based prediction of MSI. So, it is just based on the histopathology slides. The principle with all of these AI methods is that the more data you have, the more variability you have in your training data, the better your performance gets and here, you can see that we trained our MSI prediction system, on increasing numbers of tumours, increasing numbers of patients, and then we always tested it on the same cohort. And the performance went up and up and up, until at some hope plateaued, maybe at a few thousand patients. Other groups have shown similar effects that the basically the arrow here goes on the right-hand side, the arrow goes down, and the more patients you train on it. And then what you need for AI in medicine, is really not just proof of concept publications, but also large-scale validation. And this is what we did for our AI best MSI prediction system. We retrained our networks on many, many thousands of patients and then from many countries and then validated them on how their cohorts and the performance was really, high in this case. However, this is still not a clinically usable product, so in order to turn this into a diagnostic device, and of course a lot of regulatory steps need to be taken and that's basically the realm of industry which we are collaborating with in this case. But MSI in colon cancer is of course, just one example of a molecular alteration that's clinically that changes our clinical management. There is also many other mutations, and what we did in the follow-up study was to look at basically dozens of tumour-types, and dozens of molecular alterations in these tumour types, and checked which ones can we predict directly from the histopathology images. And it turned out that actually, in the majority of all genetic alterations, we cannot predict from histology, but across the board, there were like a third of all the molecular alterations we looked at that we could predict to some degree from histopathology. And in colon cancer, there were some clinically interesting candidates here, for example, TP53, KRAS and BRAF. Especially BRAF, is a clinically relevant genetic alteration in colon cancer, which is now also directly targetable in second line treatment of metastatic colon cancer. And we can directly target BRAF mutant tumours. And therefore, it's a very, very interesting candidate, also to look at. Well, from a scientific point of view, an important thing is that we train these deep learning systems with our ground truth data that is somehow available for each patient. So, we know patient A is BRAF mutant, and patient B is BRAF wild type. And then we take these ground truth labels, and we put them together with the images and we train the system. However, the predictions that we get, they are not just defined for a given patient, but they're also spatially resolved. So, for all the patients in our test set, we also, in addition to a prediction for the whole patient, we get somehow a spatially resolved prediction that tells us this area of the tumour was more likely to be BRAF mutant than the other area. And this is in fact, a very, very interesting consistent effect that we see in many publications that I think shows that these methods can not only help us to maybe automate workflows in the clinic, or make them a little more efficient, but also to identify tumour heterogeneity, to get new insights into the tumour biology. So, I think these methods can be a scientifically valuable tool in addition to potentially clinical and diagnostically relevant too. Ultimately, what we do in the clinic, however, is that we integrate a lot of different types of information, so we would never make a diagnose just, or enter treatment recommendation, just based on a single tissue slide, but of course, we looked at the patient and their medical history, genetic data, radiology images, laboratory tests, et cetera. And what we, human doctors do, is to

integrate these different types of information. And this is something that we can really do much, much better than the computer. And so maybe you'll remember that in 2012, there was this big initiative, IBM Watson in oncology, where they tried to somehow automate the whole process in where this big company, IBM, tried to automate the whole diagnostic prognostic process in oncology, essentially tried to do automatic tumour board, computer-based tumour board. And this initiative did somehow not lead to any clinically tangible benefit until now. And so, that's because it's really hard for computers to integrate different types of data. And this is something where we humans still are much better than computers, but this is also where I think the field is moving in the future, so as we have more structured data available, as we have more better methods and better hardware. We see more and more academic studies moving into this field of multimodal data and analysis and maybe in the next few years, we will also move beyond biomarkers that are just defined on a single type of image, more to towards multimodal biomarkers. And then biomarkers that somehow use AI to integrate different types of information sources. And so, what do we actually need to make this happen in the clinic? To develop these methods to push the field forward. Well, I think we need four ingredients. One of them is to have the hardware. One of them is to have the algorithms and the data. And these three things are really solvable, so hardware you can really buy any consumer hardware and for a few thousand euros, to run these state-of-the-art AI methods. The algorithms are pretty well-established and openly available on the internet, so that something that you can reuse and apply to this specific medical problem. Then of course the data is always challenged, but we have of course, public data repositories and then with a little bit of work, we can also use data from our clinical routine, but really the limiting factor, and what we need to work on, is to have an interdisciplinary bunch of people who can use these tools who can ask the right questions and implement it in real-world applications. And this is exactly what we try to do in our research group here in Aachen, Germany. And if you want to learn more about this, check out our website, www.kather.ai. And this is my interdisciplinary team. We have medical doctors, medical students, engineers, computer scientists, biologists, and it's certainly a lot of fun to collaborate with all of them. So, I want to thank my team and also thank you for your attention.