European Journal of Cancer 182 (2023) 122–131

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

Fast and label-free automated detection of microsatellite status in early colon cancer using artificial intelligence integrated infrared imaging

Klaus Gerwert [a,1], Stephanie Schörner [a,1], Frederik Großerueschkamp [a], Anna–Lena Kraeft [b], David Schuhmacher [c], Carlo Sternemann [d], Inke S. Feder [d], Sarah Wisser [d], Celine Lugnier [b], Dirk Arnold [e], Christian Teschendorf [f], Lothar Mueller [g], Nina Timmesfeld [h], Axel Mosig [c], Anke Reinacher-Schick [b,1], Andrea Tannapfel [d,*,1]

[a] *Center for Protein Diagnostics (PRODI), Deptartment of Biophysics, Ruhr University Bochum, Bochum, Germany*
[b] *Deptartment of Haematology, Oncology and Palliative Care, St. Josef-Hospital, Ruhr University Bochum, Bochum, Germany*
[c] *Center for Protein Diagnostics (PRODI), Dept. of Bioinformatics, Ruhr University Bochum, Bochum, Germany*
[d] *Institut für Pathologie, Ruhr-Universität Bochum, Bochum, Germany*
[e] *Oncology, Haematology, Palliative Care Deptartment Asklepios Tumorzentrum Hamburg AK Altona, Hamburg, Germany*
[f] *Internal Medicine, Medizinische Klinik St.-Josefs-Hospital, Dortmund, Germany*
[g] *Onkologie UnterEms Leer Emden Papenburg, Onkologische Schwerpunktpraxis Leer-Emden, Leer, Germany*
[h] *Medical Informatics, Biometry and Epidemiology, Ruhr University Bochum, Bochum, Germany*

Received 4 October 2022; received in revised form 14 December 2022; accepted 23 December 2022
Available online 9 January 2023

KEYWORDS
Microsatellite instability;
Colon cancer;
Artificial intelligence;
Deep learning;
Infrared imaging;
Label-free;
Convolutional neural networks

Abstract ***Purpose:*** Microsatellite instability (MSI) due to mismatch repair (MMR) defects accounts for 15–20% of colon cancers (CC). MSI testing is currently standard of care in CC with immunohistochemistry of the four MMR proteins representing the gold standard. Instead, label-free quantum cascade laser (QCL) based infrared (IR) imaging combined with artificial intelligence (AI) may classify MSI/microsatellite stability (MSS) in unstained tissue sections user-independently and tissue preserving.
Methods: Paraffin-embedded unstained tissue sections of early CC from patients participating in the multicentre AIO ColoPredict Plus (CPP) 2.0 registry were analysed after dividing into three groups (training, test, and validation). IR images of tissue sections using QCL-IR microscopes were classified by AI (convolutional neural networks [CNN]) using a two-step

* *Corresponding author*: Institute of Pathology, Ruhr University Bochum, Bürkle-de-la-Camp-Platz 1, Bochum, 44789, Germany.
E-mail address: andrea.tannapfel@pathologie-bochum.de (A. Tannapfel).
[1] Both authors contributed equally.

https://doi.org/10.1016/j.ejca.2022.12.026

approach. The first CNN (modified U-Net) detected areas of cancer while the second CNN (VGG-Net) classified MSI/MSS. End-points were area under receiver operating characteristic (AUROC) and area under precision recall curve (AUPRC).

Results: The cancer detection in the first step was based on 629 patients (train n = 273, test n = 138, and validation n = 218). Resulting classification AUROC was 1.0 for the validation dataset. The second step classifying MSI/MSS was performed on 547 patients (train n = 331, test n = 69, and validation n = 147) reaching AUROC and AUPRC of 0.9 and 0.74, respectively, for the validation cohort.

Conclusion: Our novel label-free digital pathology approach accurately and rapidly classifies MSI vs. MSS. The tissue sections analysed were not processed leaving the sample unmodified for subsequent analyses. Our approach demonstrates an AI-based decision support tool potentially driving improved patient stratification and precision oncology in the future.

1. Introduction

Microsatellite instability (MSI) is a distinct genomic feature caused by a defective mismatch repair system (dMMR) present in multiple cancers. It occurs in approximately 15% of colorectal cancer (CRC) cases [1]. MSI is stage-dependent, with over 20% detected in stage II and less than 5% in advanced CRC (aCRC), referring to stage IV metastatic CRC. MSI tumours are characterised by poor differentiation, mucinous component, proximal location, female, and older age [2]. MSI tumours also occur in younger patients with Lynch syndrome, the most common form of hereditary colon cancer (CC) [1]. MSI tumours are associated with a high tumour mutational burden deriving from the MSI-driven oncogenic pathway, (over-)expressing highly immunogenic neoantigens, and immune checkpoints. MSI has long been established as a positive prognostic parameter in stage II CRC and has now become an important predictive marker for the efficacy of immune checkpoint inhibition in multiple cancer types including CRC [3,4], setting the new standard of care in this molecular patient subgroup. Immune checkpoint inhibitors are currently also being evaluated in randomised controlled trials in localised CRC both in adjuvant and neoadjuvant settings [1]. Testing of MSI is therefore recommended in CRC of all stages [5].

Two standard reference methods, immunohistochemistry (IHC) and fragment length analysis by polymerase chain reaction (PCR), are used to screen for MSI. IHC analysis of the MMR proteins MLH1, MSH2, MSH6, and PMS2 is the gold standard in clinical routine [6]. Next-generation sequencing is an alternative test for MSI with approval by the Federal Drug Administration for MSI determination in aCRC. Considering its crucial positive predictive value for checkpoint inhibitor therapy, misdiagnosis of MSI, which has been reported in up to 10% for IHC, may result in substantial disadvantages for the individual patient [7]. Furthermore, immunohistochemical staining leaves tissue unusable for further molecular analysis.

With the advances in computing power and the possibility of utilising big data, artificial intelligence (AI)-assisted digital pathology has gained significant attraction [8–11]. MSI was detected from haematoxylin–eosin (H&E)-stained sections applying a deep learning approach requiring very large patient cohorts for training, showing impressive validation [12]. However, all methods for MSI detection so far depend on variously modified tissue – rendering the sample inaccessible for further molecular characterisation.

Infrared (IR) imaging is an emerging label-free technology for unstained tissue sections. A quantum cascade laser (QCL) based IR microscope records spatially resolved infrared spectra providing 427 molecular-specific channels. Previously, precise, operator-independent identification of cancer, subtypes, grading, and molecular alterations in unstained tissue sections has been reported [13–15]. Feasibility of IR-based MSI detection using a machine learning approach has previously been demonstrated by our group using a random forest classifier trained on 40 patients and tested on 60 patients [16]. Although a very good area under receiver operating characteristic curve (AUROC) could be achieved on the small collective, an internal validation on a larger cohort from the multicentre AIO (*Arbeitsgemeinschaft internistische Onkologie*) ColoPredict Plus (CPP) 2.0 registry showed that the robustness and generalisability was not given. To overcome these limitations, the presented study aimed to test and validate an IR imaging AI-based approach to classify MSI vs. microsatellite stability (MSS) in early CC tissue samples from the AIO CPP 2.0 registry. This approach now allows us for the first time a robust detection of MSI in CC by IR imaging.

2. Material and methods

2.1. Patients

The multicentre AIO CPP 2.0 registry trial (DRKS00004305) captures clinical data and tissue samples of patients with early CC in 188 trial centres throughout Germany [17]. All study participants provided informed consent, and the study protocol was approved by the institutional review board (IRB) of the Ruhr-University Bochum (registration number: 4453-12 & 20-6830) and by local IRBs of participating sites. The sample study is based on feasibility studies conducted with IR imaging and a statistical evaluation for the expected length of the 95% confidence interval (CI) for the sensitivity and specificity using the Agresti–Coull method [16,18]. The study design and report of results followed the STAndards for Reporting Diagnostic accuracy studies (STARD) guidelines as described in Supplementary Table 1.

According to the study protocol, prerequisites for the selection of patients were age above 18, CC stage I, II, and III, written consent, and known microsatellite status. All samples are surgical specimens, and MSI status was assessed centrally using immunohistochemistry in combination with PCR-based fragment length analysis as described recently [19]. The selection aimed to balance parameters of age, sex, localisation, BRAF, and RAS mutational status. Randomly assigned training (optimisation of the neural network) and testing (model selection) samples were chosen with a greater proportion of MSI cases to stabilise the training process with a more balanced dataset. Samples for the validation cohort were randomly selected and therefore show a natural occurrence of MSI cases. The full study flowchart is illustrated in Supplementary Figure 1.

2.2. Sample processing

Tissue samples were submitted from participating sites and collaborating institutes of pathology as formalin-fixed, paraffin-embedded (FFPE) samples and therefore passed through different, locally established, pre-analytic procedures. For IR imaging measurements using the QCL-based IR microscope, 7 µm tissue sections were mounted on IR-suitable polyethylene terephthalate frame slides and measured as described previously [20]. Subsequently, samples were H&E stained and digitised using an Olympus slide scanner (Olympus VS120).

2.3. Infrared imaging

The QCL-based microscope was used in the standard transmission configuration (Supplementary Figure 2) [15]. The light sources are QCLs implemented in the platform Spero QT (Daylight Solutions, San Diego, CA, USA). This QCL platform analyses tissue sections within 30 min including AI-based classification. Detailed measurement parameters are provided in Supplementary Table 2, and study workflow is illustrated in Supplementary Figure 3.

2.4. Deep learning algorithms

For classification of the IR images, we used convolutional neural networks (CNN) in a two-step setup: during step 1, cancer areas were detected within the IR images, while in step 2, MSI/MSS was classified within the detected tumour tissue (Fig. 1).

The step 1 CNN is an in-house developed weakly supervised extension of a U-net architecture, named comparative segmentation network (CompSegNet) [20,21], resulting in pixel-precise *cancer detection*. The CompSegNet was trained using subregions from the whole slide IR images that were assigned by an experienced pathologist (A.T.). The CompSegNet training was conducted on these manually selected and binary labelled subregions of size 256×256 pixels (px) (Supplementary Info 1). Each region that was labelled as cancer contained a minimum area of 20% cancer. Cancer areas identified in step 1 were then used for the step 2 CNN – a VGG classification network (see Fig. 1) – for the training of the *MSI/MSS classifier* [22]. The technical parameters for the deep learning workflow are presented in Supplementary Table 3. For the application of the CNNs on whole slide IR images, the detection/classification was performed for each subregion in the whole slide IR image and merged for a whole slide result.

2.5. Statistical analysis

The statistical evaluation was carried out in two levels: on the single input subregions and on the result for the whole slides (averaged for all subregions) on patient level. The subregion-based evaluation serves as assessment of the technical AI approach, whereas the whole slide evaluation reflects the performance relevant for clinical application (details in Supplementary Info 2).

The performance of the classification process was evaluated using AUROC and area under precision recall curve (AUPRC) defining MSI samples as positive and MSS samples as negative. Here, 95% robust pointwise CI for AUROC and AUPRC were generated by bootstrapping. In addition, sensitivity, specificity, accuracy, and F1 score (definition in Supplementary Info 4) were determined by selecting a cut-off value for the prediction (range: 0 to 1) balancing sensitivity and specificity on the test cohort. Additionally, dice score was determined for step 1 cancer detection using 25 randomly selected samples from the validation cohort.

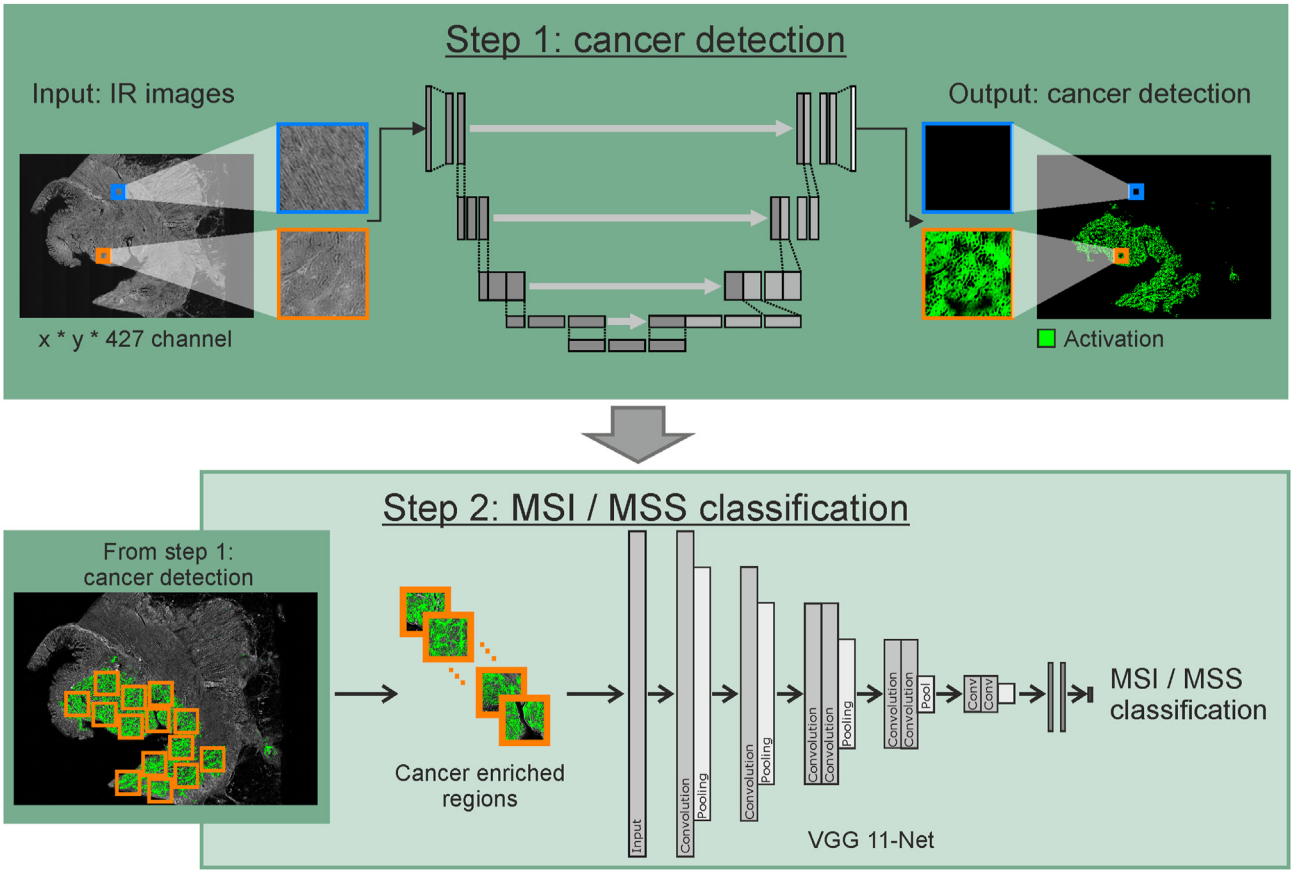


Fig. 1. **The rapid and fully automated two-step approach for MSI/MSS classification via IR imaging.** The first step uses a U-net based architecture (CompSegNet) as a weakly supervised approach for the cancer detection. The input is manually selected subregions of size 256×256 pixels, here shown in orange (cancer) and blue (non-cancer), with 427 optical IR channels. The output detection is further used for the selection of cancer enriched subregions as input for the VGG 11 Neural Network, which performs the MSI/MSS classification. (IR: infrared; MSI: microsatellite instability; MSS: microsatellite stability). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article).

3. Results

3.1. Patients

As shown, 633 patients were selected from the AIO CPP 2.0. Four samples were excluded before the analysis after quality check of the IR measurements. The step 1 *cancer detection* cohort therefore consisted of 629 patients as shown in Supplementary Table 4. A total of 10,942 subregions were selected manually from the IR measurements (training: 4955 – test: 2348 – validation: 3639) representing cancer and cancer-free regions and used for training, model selection, and statistical evaluation.

The step 2 *MSI/MSS classifier* sub-cohort included 547 patients (patients with tissue sections containing cancerous regions from the abovementioned cohort) as presented in Table 1 and Supplementary Figure 1. Patients with only normal tissue sections (n = 80) or samples with insufficient cancer material for further analysis (n = 2) were excluded. The characteristics are equally distributed except MSI for methodological reasons (see *Material and methods* section).

3.2. Label-free cancer detection

The step 1 CNN achieved an AUROC of 1.0 (95% CI 1.000) for whole slide classification, thus, forming the basis for the *MSI/MSS classifier*. [20] All parameters (AUROC, AUPRC, F1 score, sensitivity, specificity, and accuracy) achieved a value of 1.0 on the whole slide level for cancer classification (see Fig. 2A and B), and the *cancer detector* reaches a dice score of 0.5. As shown in Fig. 2, the detection of cancer and cancer-associated tissue performs very precisely. Whole slide images illustrating the step 1 *cancer detection* are shown in Fig. 2C with pixel-precise detection illustrated in green. A cancer-free sample (NT = no tumour) is shown for comparison.

3.3. MSI/MSS classification

The step 2 *MSI/MSS classifier* was trained on cancer subregions detected by step 1.1500 epochs (training curves in Supplementary Figure 4) were trained, creating an equal number of classifier models. The best

Table 1
Patient cohort for MSI/MSS classification.

		Total	Training		Test		Validation	
		MSI + MSS	MSI	MSS	MSI	MSS	MSI	MSS
	N (% of train/test/vali)[a]	**547**	142 (*43*)	189 (*57*)	30 (*43*)	39 (*57*)	26 (*18*)	121 (*82*)
Sex	f/m in %[b]	***50/50***	*64/36*	*40/60*	*67/33*	*31/69*	*65/35*	*50/50*
Age	Mean ± SD (years)	**69 ± 12**	71 ± 14	68 ± 12	73 ± 12	70 ± 11	73 ± 10	66 ± 12
Location	left (%[b])	**211 (*39*)**	30 (*21*)	98 (*52*)	6 (*20*)	21 (*54*)	3 (*12*)	53 (*44*)
	Right (%[b])	**331 (*61*)**	112 (*79*)	90 (*48*)	24 (*80*)	18 (*46*)	23 (*88*)	64 (*53*)
	overlap (%[b])	**5 (*1*)**	0 (*0*)	1 (*0*)	0 (*0*)	0 (*0*)	0 (*0*)	4 (*3*)
UICC	I (%[b])	**13 (*2*)**	9 (*6*)	1 (*0*)	2 (*6*)	0 (*0*)	1 (*4*)	0 (*0*)
	II (%[b])	**152 (*28*)**	64 (*45*)	37 (*20*)	14 (*47*)	8 (*20*)	16 (*61*)	13 (*11*)
	III (%[b])	**382 (*70*)**	69 (*49*)	151 (*80*)	14 (*47*)	31 (*80*)	9 (*35*)	108 (*89*)
Grading	G1 (%[b])	**21 (*4*)**	8 (*6*)	3 (*2*)	2 (*7*)	1 (*3*)	3 (*11*)	4 (*3*)
	G2 (%[b])	**343 (*63*)**	59 (*42*)	144 (*76*)	11 (*37*)	27 (*69*)	13 (*50*)	89 (*74*)
	G3 (%[b])	**160 (*29*)**	67 (*47*)	36 (*19*)	16 (*53*)	6 (*15*)	9 (*35*)	26 (*21*)
	G4 (%[b])	**2 (*0*)**	2 (*1*)	0 (*0*)	0 (*0*)	0 (*0*)	0 (*0*)	0 (*0*)
	unknown (%[b])	**21 (*4*)**	6 (*4*)	6 (*3*)	1 (*3*)	5 (*13*)	1 (*4*)	2 (*2*)
KRAS	wt (%[b])	**372 (*68*)**	119 (*84*)	113 (*60*)	24 (*80*)	24 (*62*)	16 (*62*)	76 (*63*)
	mt (%[b])	**175 (*32*)**	23 (*16*)	76 (*40*)	6 (*20*)	15 (*38*)	10 (*38*)	45 (*37*)
BRAF	wt (%[b])	**423 (*77*)**	72 (*51*)	171 (*90*)	17 (*57*)	36 (*92*)	15 (*58*)	112 (*93*)
	mt (%[b])	**124 (*23*)**	70 (*49*)	18 (*10*)	13 (*43*)	3 (*8*)	11 (*42*)	9 (*7*)
NRAS	wt (%[b])	**534 (*98*)**	139 (*98*)	182 (*96*)	29 (*97*)	39 (*100*)	26 (*100*)	119 (*98*)
	mt (%[b])	**13 (*2*)**	3 (*2*)	7 (*4*)	1 (*3*)	0 (*0*)	0 (*0*)	2 (*2*)

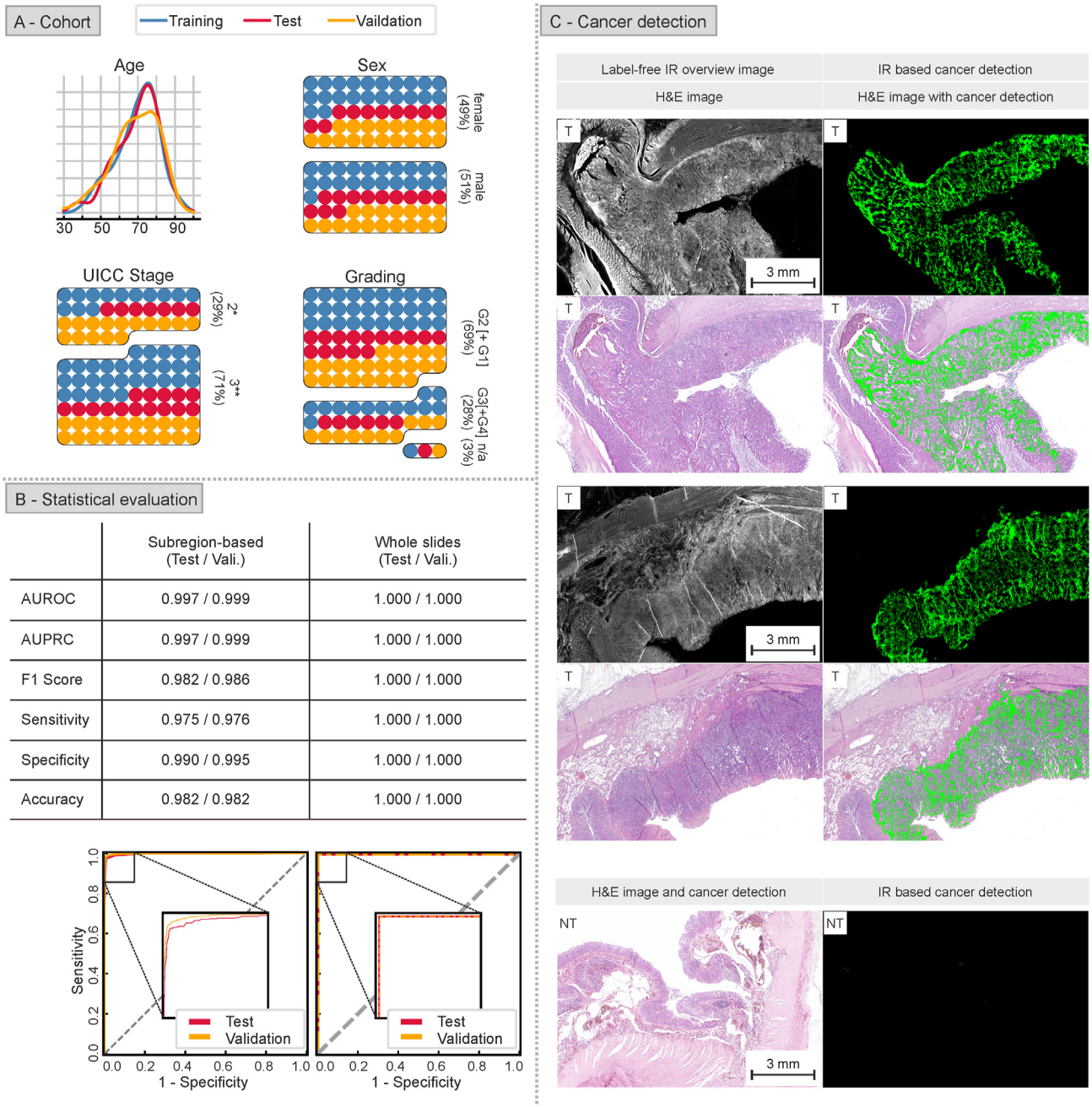
MSI, microsatellite instability; MSS, microsatellite stability; SD, standard deviation. The numbers for the total cohort are printed in bold.
[a] Percentage calculated for training (train)/test/validation (vali) cohorts.
[b] Percentage calculated column-wise and are printed in italics.

performing model from epoch 1000 was selected based on the AUROC value of the test set. Fig. 3 summarises the statistical results and illustrates the application on whole slide images. The subregion-based evaluation performance was comparable to the whole slide evaluation, with AUROC values of 0.87 (95% CI 0.78–0.97) and 0.90 (95% CI 0.83–0.97) and AUPRC values of 0.89 (95% CI 0.80–0.98) and 0.74 (95% CI 0.59–0.89) obtained for test and validation, respectively, in the whole slide setting. Classification on the validation cohort had 85% sensitivity and 84% specificity. ROC curves, PR curves, and statistical parameters characterising the three cohorts are provided in Supplementary Figure 5 and 6, including a detailed analysis of the 95% CI. Steps 1 and 2 are summarised in Fig. 3C illustrating the process of *cancer detection* and subsequent *MSI/MSS classification* within the cancer regions. For the whole slide results of step 2, green represents high predictive values for MSI, whereas purple represents high predictive values for MSS.

4. Discussion

At present, two standard reference methods, IHC and PCR, are recommended for the detection of MSI in CRC. IHC is considered the gold standard in routine practice followed by other molecular approaches such as PCR when IHC is equivocal. Next-generation sequencing has been reported to exhibit high concordance with IHC and PCR and has recently been approved by the Federal Drug Administration in aCRC [24]. All techniques listed above use formalin-fixed and stained tumour tissue, thereby limiting its availability for further tissue-based analyses. MSI has long been known as a positive prognostic marker in early CRC. Most guidelines now recommend testing for MSI in patients with low-risk stage II cancer to avoid the use of ineffective fluoropyrimidines as adjuvant treatment in this setting [5,30]. Recently, MSI has emerged as a crucial predictive biomarker for the efficacy of immune checkpoint inhibition (e.g. pembrolizumab) which is now regarded standard of care in MSI aCRC and is also evaluated in early CRC in clinical trials [3]. The diagnosis of MSI must therefore be highly reliable. Presently, misdiagnosis by IHC has been reported to be as high as 10% resulting in a substantial disadvantage for the individual patient [7]. With the rapid IR imaging technique (30 min turn-around time), we were able to reliably distinguish MSI from MSS in the tissue of patients with early CC achieving high sensitivity and specificity, comparable to the present gold standard (IHC) [7,25]. With this novel technique tissue samples are kept available for further molecular assessment.

We developed and validated a novel automated and label-free digital pathology approach to distinguish MSI from MSS reliably and rapidly in CC tissue samples. FFPE samples from a total of 547 patients from the German AIO CPP registry trial were analysed with centrally determined MSI/MSS status. These tissue samples were initially processed by local standards at different institutes of pathology by investigators, thus



	Subregion-based (Test / Vali.)	Whole slides (Test / Vali.)
AUROC	0.997 / 0.999	1.000 / 1.000
AUPRC	0.997 / 0.999	1.000 / 1.000
F1 Score	0.982 / 0.986	1.000 / 1.000
Sensitivity	0.975 / 0.976	1.000 / 1.000
Specificity	0.990 / 0.995	1.000 / 1.000
Accuracy	0.982 / 0.982	1.000 / 1.000

Fig. 2. **Results for the step 1 cancer detection.** (A) Illustrates the patient cohort and the distribution between training, test, and validation sets. The table in (B) lists the determined statistical parameters for the test and validation cohort, both for the analysis with the predictions on the pre-selected subregions and the whole slide images. Below each column the corresponding ROC curves are shown. (C) Shows the results of the whole slide prediction for three examples from the validation cohort: the upper two examples show cancer sections (T) with the corresponding IR overview image and the H&E image, and the NT labelled image shows a non-tumour (NT) section. The H&E images have been registered onto the IR images according to [23] (IR: infrared).

increasing variation in the dataset and therefore adding to the robustness of the method. This includes slight variations in the thickness of the tissue sections, which did not affect the classification results. In the first step, the approach was able to classify cancer with an AUROC of 1.0 (95% CI 1.00–1.00) and a dice score of 0.5 in the validation cohort of 218 patients. MSI vs. MSS could be correctly classified with an AUROC value of 0.9 (95% CI 0.83–0.97) and an AUPRC value of 0.74 (95% CI 0.59–0.89) in a validation cohort of 147 patients. The AUROC in the validation cohort increased compared to the test cohort, which may be due to a limited number of MSI cases in the validation cohort following the natural occurrence of MSI. Overall, the

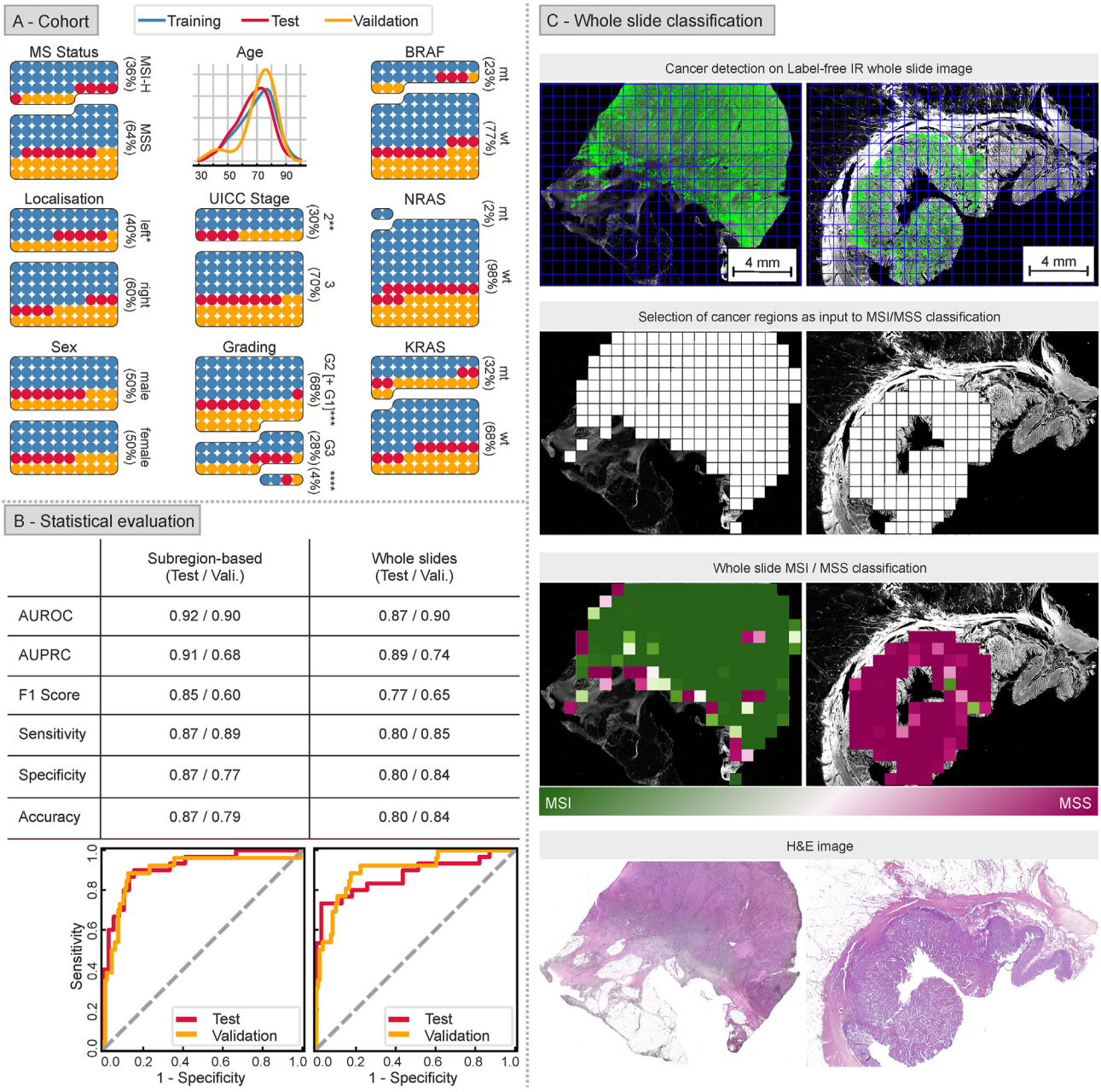


Fig. 3. **Results for the step 2 MSI/MSS classification.** (A) Illustrates the patient cohort and the distribution between training, test, and validation cohort [∗∗including overlap, ∗∗including 2% with stage I, ∗∗∗including 4% G1, ∗∗∗∗including two patients with G4, rest unknown]. The table in (B) lists the statistical parameters for the test and validation cohort, together with the ROC curves for the test and validation cohort on patient-level. Samples with predictive vales above the cut-off of 0.48 (subregion-based) and 0.39 (whole slide) were evaluated as MSI cases. Two examples for a whole slide classification from the validation cohort are shown in (C): The classification process is presented from top to bottom and starts with the cancer detection (light green) by a moving window (raster). The cancer subregions are selected and used as input for the MSI/MSS classifier. For the final classification, green areas indicate a strong predictive correlation to MSS whereas pink areas indicate a strong predictive correlation to MSI. As comparison, the H&E image is presented additionally at the bottom. (IR: infrared; MSI: microsatellite instability; MSS: microsatellite stability). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article).

modified architecture of the VGG network exhibited a stable training behaviour. Our approach could thus also be reliably used in a multicentre setting and has the potential to be transferred to institutions and caregivers globally in the future.

The label-free digital pathology approach described here can detect MSI vs. MSS with sensitivity and specificity comparable to the approach, which has been previously described on H&E-stained images [12,26,27]. Compared to the results from Yamashita *et al.*, the

label-free approach achieves similar accuracy without the need for staining or other processing of the sample. The H&E-based approaches by Echle *et al.* showed best results on a 5500-patient cohort with an AUROC of 0.92. In their study, training sets consisting of 500 patients reached an average AUROC of only 0.82. The label-free approach described here achieved an AUROC of 0.9 (95% CI 0.83–0.97) on an even smaller 331-patient training dataset, emphasising the information density that IR imaging can retrieve in unstained samples. This is an advantage of IR imaging regarding smaller molecular or clinical subgroups as well as potentially detecting novel predictive biomarkers, e.g., from randomised trials with limited numbers of participants.

This label-free approach enables the detection of areas of interest in unstained tumour material, making, beyond the immediate clinical application, tumour heterogeneity analysis, elucidation of underlying pathways, and provision of enriched material for multi-omics approaches possible [13,28,29]. In contrast to other stain-based methods, the label-free IR approach allows pixel-precise dissection on the same sample slide the detection has been performed on.

The determination of MSI is also clinically relevant in multiple other solid tumours such as gastric or endometrial cancer. Various other molecular alterations are of impact on clinical management (RAS, BRAF, HER2), which have different prevalence in MSI vs. MSS cancers. Therefore, this reliable *MSI/MSS classification* method may also serve as a screening method to identify patients for subsequent molecular analyses.

5. Conclusion

This label-free digital pathology approach for *MSI/MSS classification* is a novel and valid tool for stratification in precision oncology. The analysis of unstained tissue sections moreover allows targeted, subsequent molecular analysis on microdissectable, predefined areas of interest. Overall, this method has the potential to become an applicable and valuable diagnostic tool for stratification beyond the scope of known biomarkers for tumour-agnostic assessments.

Credit author statement

Klaus Gerwert: Conceptualisation, Methodology, Investigation, Resources, Writing – Review and editing, Supervision, Project administration, Funding acquisition. **Stephanie Schörner:** Conceptualisation, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – Original draft, Visualisation. **Frederik Großerüschkamp:** Conceptualisation, Methodology, Investigation, Writing – Original draft, Supervision, Project administration, Funding acquisition. **Anna–Lena Kraeft:** Writing – Review and editing, Project administration. **David Schuhmacher:** Methodology, Software. **Carlo Sternemann:** Data curation. **Inke Feder:** Validation, Formal analysis, Data curation. **Sarah Wisser:** Data curation. **Celine Lugnier:** Writing – Review and editing. **Dirk Arnold:** Resources, Writing – Review and editing. **Christian Teschendorf:** Resources, Writing – Review and editing. **Lothar Mueller:** Resources, Writing – Review and editing. **Nina Timmesfeld:** Conceptualisation, Validation, Formal analysis, Writing – Review and editing. **Axel Mosig:** Methodology, Software, Writing – Review and editing, Supervision, Funding acquisition. **Anke Reinacher-Schick:** Conceptualisation, Resources, Writing – Review and editing, Supervision, Project administration, Funding acquisition. **Andrea Tannapfel:** Conceptualisation, Resources, Writing – Review and editing, Supervision, Project administration, Funding acquisition.

Conflict of interest statement

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: **Dirk Arnold** reports Consultation/Advisory role from AstraZeneca, Bristol Myers Squibb, Merck Sharp & Dohme, Sanofi (Genzyme), Bayer, Terumo, Roche, Hexal, Samsung Bioepis, Pierre Fabre Pharma, CRA International, Ketchum, IQVIA; he reports speaker's engagement from AstraZeneca, Bristol Myers Squibb, Merck Sharp & Dohme, Sanofi (Genzyme), Bayer, Eli Lilly, Terumo, Boston Scientific, Amgen, Roche, Ipsen, Merck (Serono), Samsung Bioepis, Pierre Fabre Pharma, Servier, PharmaCept, PRMA Consulting, Tactics MD LLC, WebMD Health Corp, From Research to Practice, Aptitude Health, art tempi media, ACE Oncology, Imedex, streamitup Germany, MedAhead (Austria), Clinical Care Options (CCO); he reports Local PI for Bristol Myers Squibb, Pierre Fabre Pharma and Coordinating PI of OncoLytics; he reports grant funding from AbbVie; he reports DSMB chair of Sanofi (Genzyme); he reports Steering Committee Member of Roche. **Anke Reinacher-Schick** reports Consulting or Advisory Role from Amgen, Merck Serono, Roche Pharma AG, Bristol-Myers Squibb, MSD, AstraZeneca, Pierre Fabre; she reports Honoria from Amgen, Merck Serono, Roche Pharma AG, Bristol-Meyers Squibb, MSD, MCI Group, AstraZeneca; she reports research funding for her institution from Celgene (Inst), Roche Pharma AG (Inst), Ipsen, Amgen, Alexion Pharmaceuticals, AstraZeneca, Lilly, Servier, AIO-Studien, Rafael Pharmaceuticals, ERYTECH Pharma, BioNTech; she reports Travel/Accommodations/Expenses from Roche. **Andrea Tannapfel** reports Honoria from Amgen, Roche, Merck Senoro, Bristol-Myers Squibb, MSD, AstraZeneca; she reports research funding from

Roche, Amgen, BioNTech. All remaining authors have declared no conflicts of interest.

Funding

This work was supported by the Ministry of Culture and Science (MKW) of the State of North-Rhine Westphalia, Germany [grant number: 111.08.03.05–133974] and by Roche Pharma AG to the Colo-Predict Plus 2.0 registry (no grant number). Part of this research was conducted within the Slide2Mol project funded by the Computational Life Sciences program of the German Federal Ministry of Education and Research, grant number 031L0264.

Data sharing

The spectral and medical data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements

The authors thank the patients participating in the CPP registry study and their families for allowing us to realise this research work. The authors further thank the staff involved at the CPP study sites all over Germany for their contribution. The authors thank Dr. Claus Küpper and Melina Helfrich for their support in data acquisition, Lidia Janota for her assistance in sample preparation, and Dr. Oliver Overheu for his valuable help in describing the cohort.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejca.2022.12.026.

References

[1] Kim JC, Bodmer WF. Genomic landscape of colorectal carcinogenesis. J Cancer Res Clin Oncol 2022;148:533–45.
[2] Dienstmann R, Mason MJ, Sinicrope FA, et al. Prediction of overall survival in stage II and III colon cancer beyond TNM system: a retrospective, pooled biomarker study. Ann Oncol 2017;28:1023–31.
[3] André T, Shiu K-K, Kim TW, et al. Pembrolizumab in microsatellite-instability-high advanced colorectal cancer. N Engl J Med 2020;383:2207–18.
[4] Cercek A, Lumish M, Sinopoli J, et al. PD-1 blockade in mismatch repair–deficient, locally advanced rectal cancer. N Engl J Med 2022;386:2363–76.
[5] Argilés G, Tabernero J, Labianca R, et al. Localised colon cancer: ESMO clinical practice guidelines for diagnosis, treatment, and follow-up. Ann Oncol 2020;31:1291–305.
[6] Boland CR, Thibodeau SN, Hamilton SR, et al. A national cancer institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 1998:5248–57.
[7] Cohen R, Hain E, Buhard O, et al. Association of primary resistance to immune checkpoint inhibitors in metastatic colorectal cancer with misdiagnosis of microsatellite instability or mismatch repair deficiency status. JAMA Oncol 2019;5:551–5.
[8] van der Laak J, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. Nat Med 2021;27:775–84.
[9] Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. Lancet Oncol 2019;20:e253–61.
[10] Naik N, Madani A, Esteva A, et al. Deep learning-enabled breast cancer hormonal receptor status determination from base-level H&E stains. Nat Commun 2020;11:5727.
[11] Chen M, Zhang B, Topatana W, et al. Classification and mutation prediction based on histopathology H&E images in liver cancer using deep learning. NPJ Precis Oncol 2020;4:14.
[12] Echle A, Grabsch HI, Quirke P, et al. Clinical-grade detection of microsatellite instability in colorectal tumors by deep learning. Gastroenterol 2020;159:1406–1416.e11.
[13] Goertzen N, Pappesch R, Fassunke J, et al. Quantum cascade laser-based infrared imaging as a label-free and automated approach to determine mutations in lung adenocarcinoma. Am J Pathol 2021;191:1269–80.
[14] Kuepper C, Großerueschkamp F, Kallenbach-Thieltges A, et al. Label-free classification of colon cancer grading using infrared spectral histopathology. Faraday Discuss 2016;187:105–18.
[15] Kuepper C, Kallenbach-Thieltges A, Juette H, et al. Quantum cascade laser-based infrared microscopy for label-free and automated cancer classification in tissue sections. Sci Rep 2018;8:7717.
[16] Kallenbach-Thieltges A, Großerueschkamp F, Jütte H, et al. Label-free, automated classification of microsatellite status in colorectal cancer by infrared imaging. Sci Rep 2020;10:10161.
[17] Noepel-Duennebacke S, Juette H, Feder IS, et al. Die hohe Mikrosatelliteninstabilität (MSI-H) ist assoziiert mit klinischen und molekularen Charakteristika sowie einem verlängerten Überleben in frühen Kolonkarzinomen: real-world Daten des molekularen Registers der AIO Colopredict Plus. Z Gastroenterol 2020;58:533–41.
[18] Agresti A, Coull BA. Approximate is better than "exact" for interval estimation of binomial proportions. Am Statistician 1998;52:119–26.
[19] Noepel-Duennebacke S, Juette H, Schulmann K, et al. Microsatellite instability (MSI-H) is associated with a high immunoscore but not with PD-L1 expression or increased survival in patients (pts.) with metastatic colorectal cancer (mCRC) treated with oxaliplatin (ox) and fluoropyrimidine (FP) with and without bevacizumab (bev): a pooled analysis of the AIO KRK 0207 and RO91 trials. J Cancer Res Clin Oncol 2021;147:3063–72.
[20] Schuhmacher D, Schörner S, Küpper C, et al. A framework for falsifiable explanations of machine learning models with an application in computational pathology. Med Image Anal 2022;82:102594.
[21] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. Medical image computing and computer-assisted intervention – MICCAI 2015. Lecture notes in computer science. Cham: Springer International Publishing; 2015. p. 234–41.
[22] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint 2014.
[23] Trukhan S, Tafintseva V, Tøndel K, et al. Grayscale representation of infrared microscopy images by extended multiplicative signal correction for registration with histological images. J Biophot 2020;13:e201960223.
[24] Ratovomanana T, Cohen R, Svrcek M, et al. Performance of next-generation sequencing for the detection of microsatellite instability in colorectal cancer with deficient DNA mismatch repair. Gastroenterol 2021;161:814–826.e7.
[25] Svrcek M, Lascols O, Cohen R, et al. MSI/MMR-deficient tumor diagnosis: which standard for screening

and for diagnosis? Diagnostic modalities for the colon and other sites: differences between tumors. Bull Cancer 2019;106:119–28.

[26] Yamashita R, Long J, Longacre T, et al. Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. Lancet Oncol 2021;22:132–41.

[27] Kather JN, Pearson AT, Halama N, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nat Med 2019;25:1054–6.

[28] Witzke KE, Großerueschkamp F, Jütte H, et al. Integrated Fourier transform infrared imaging and proteomics for identification of a candidate histochemical biomarker in bladder cancer. Am J Pathol 2019;189:619–31.

[29] Großerueschkamp F, Bracht T, Diehl HC, et al. Spatial and molecular resolution of diffuse malignant mesothelioma heterogeneity by integrating label-free FTIR imaging, laser capture microdissection and proteomics. Sci Rep 2017;7:44829.

[30] Baxter NN, Kennedy EB, Bergsland E, et al. Adjuvant therapy for stage II colon cancer: ASCO Guideline Update. J Clin Orthod 2022;40:892–910.