

# Joint Imaging Platform for Federated Clinical Data Analytics

Jonas Scherer, MS<sup>1,3</sup>; Marco Nolden, PhD<sup>1,3,4</sup>; Jens Kleesiek, MD, PhD<sup>3,5</sup>; Jasmin Metzger, DiplInformMed<sup>1,3</sup>; Klaus Kades, MSc<sup>1,3</sup>; Verena Schneider<sup>3,5</sup>; Michael Bach, PhD<sup>3,5</sup>; Oliver Sedlaczek, MD<sup>3,5,6</sup>; Andreas M. Bucher, MD<sup>3,7</sup>; Thomas J. Vogl, MD<sup>3,7</sup>; Frank Grünwald, MD<sup>3,8</sup>; Jens-Peter Kühn, MD<sup>3,9</sup>; Ralf-Thorsten Hoffmann, MD<sup>3,9</sup>; Jörg Kotzerke, MD<sup>3,10</sup>; Oliver Bethge, DiplIng<sup>3,11</sup>; Lars Schimmöller, MD<sup>3,11</sup>; Gerald Antoch, MD<sup>3,11</sup>; Hans-Wilhelm Müller, MD<sup>3,12</sup>; Andreas Daul, DiplPhys<sup>3,13</sup>; Konstantin Nikolaou, MD, MBA<sup>3,13</sup>; Christian la Fougère, MD<sup>3,14</sup>; Wolfgang G. Kunz, MD<sup>3,15</sup>; Michael Ingrisich, PhD<sup>3,15</sup>; Balthasar Schachtner, PhD<sup>3,15,16</sup>; Jens Ricke, MD<sup>3,15</sup>; Peter Bartenstein, MD<sup>3,17</sup>; Felix Nensa, MD<sup>3,18</sup>; Alexander Radbruch, MD, JD<sup>3,18</sup>; Lale Umutlu, MD<sup>3,18</sup>; Michael Forsting, PhD<sup>3,18</sup>; Robert Seifert, MD<sup>3,19</sup>; Ken Herrmann, MD, MBA<sup>3,19</sup>; Philipp Mayer, MD<sup>3,6</sup>; Hans-Ulrich Kauczor, MD, PhD<sup>3,6,16</sup>; Tobias Penzkofer, MD<sup>3,20</sup>; Bernd Hamm, MD<sup>3,20</sup>; Winfried Brenner, PhD<sup>3,21</sup>; Roman Kloeckner, MD<sup>3,22</sup>; Christoph Düber, MD<sup>3,22</sup>; Mathias Schreckenberger, MD<sup>3,23</sup>; Rickmer Braren, MD<sup>3,24</sup>; Georgios Kaissis, MD, MHBA<sup>3,4,24,25</sup>; Marcus Makowski, MD<sup>3,24</sup>; Matthias Eiber, MD<sup>3,26</sup>; Andrei Gafita, MD<sup>3,26</sup>; Rupert Trager, MSc<sup>3,26</sup>; Wolfgang A. Weber, MD<sup>3,26</sup>; Jakob Neubauer, MD<sup>3,27</sup>; Marco Reiser, PhD<sup>3,27</sup>; Michael Bock, PhD<sup>3,27</sup>; Fabian Bamberg, MD, MPH<sup>3,27</sup>; Jürgen Hennig, PhD<sup>3,27</sup>; Philipp Tobias Meyer, MD, PhD<sup>3,28</sup>; Juri Ruf, MD<sup>3,28</sup>; Uwe Haberkorn, MD<sup>3,29</sup>; Stefan O. Schoenberg, MD<sup>3,30</sup>; Tristan Kuder, PhD<sup>3,31</sup>; Peter Neher, PhD<sup>1,3</sup>; Ralf Floca, PhD<sup>1,3,4</sup>; Heinz-Peter Schlemmer, MD, PhD<sup>2,3,5</sup>; and Klaus Maier-Hein, PhD<sup>1,4</sup>

**PURPOSE** Image analysis is one of the most promising applications of artificial intelligence (AI) in health care, potentially improving prediction, diagnosis, and treatment of diseases. Although scientific advances in this area critically depend on the accessibility of large-volume and high-quality data, sharing data between institutions faces various ethical and legal constraints as well as organizational and technical obstacles.

**METHODS** The Joint Imaging Platform (JIP) of the German Cancer Consortium (DKTK) addresses these issues by providing federated data analysis technology in a secure and compliant way. Using the JIP, medical image data remain in the originator institutions, but analysis and AI algorithms are shared and jointly used. Common standards and interfaces to local systems ensure permanent data sovereignty of participating institutions.

**RESULTS** The JIP is established in the radiology and nuclear medicine departments of 10 university hospitals in Germany (DKTK partner sites). In multiple complementary use cases, we show that the platform fulfills all relevant requirements to serve as a foundation for multicenter medical imaging trials and research on large cohorts, including the harmonization and integration of data, interactive analysis, automatic analysis, federated machine learning, and extensibility and maintenance processes, which are elementary for the sustainability of such a platform.

**CONCLUSION** The results demonstrate the feasibility of using the JIP as a federated data analytics platform in heterogeneous clinical information technology and software landscapes, solving an important bottleneck for the application of AI to large-scale clinical imaging data.

JCO Clin Cancer Inform 4:1027-1038. © 2020 by American Society of Clinical Oncology

Licensed under the Creative Commons Attribution 4.0 License 

## ASSOCIATED CONTENT

### Appendix

#### Data Supplement

Author affiliations and support information (if applicable) appear at the end of this article.

### Accepted on

September 17, 2020

and published at

[ascopubs.org/journal/](https://ascopubs.org/journal)

cci on November 9,

2020: DOI <https://doi.org/10.1200/CCI.20.00045>

00045

## INTRODUCTION

Medical imaging plays an essential role in nearly all aspects of high-quality cancer care, from preventive measures, including screening and early detection, through diagnosis, treatment planning and monitoring, and follow-up. Most patients with cancer undergo repeated imaging during the course of their treatment. In combination with clinical and laboratory data, quantitative imaging biomarkers are of fundamental importance to improve standardized therapy monitoring in clinical multicenter studies.<sup>1-7</sup>

Medical images are more than pictures; they are data characterizing the patient.<sup>8</sup> As such, they are rightly subject to strict data protection, as well as ethical

and moral requirements for scientific secondary use, which, in turn, can impede the exchange of biomedical imaging data across clinical sites. Current research strongly indicates that data anonymization is not only difficult to perform but also generally ineffective in practice.<sup>9-11</sup> De-identification or anonymization methods considered safe today might potentially fail in the future.<sup>9</sup> Data ownership, insufficient personal incentives for data collectors to share their data, and remaining technical challenges present further hurdles to data sharing in the medical research domain.<sup>12</sup> In addition, the clinical landscape is composed of heterogeneous information technology (IT) systems as well as different scanners and acquisition parameters, making joint projects and data

## CONTEXT

### Key Objective

To create a digital infrastructure for federated artificial intelligence (AI)–based medical image analysis with the goal of facilitating and enabling multicenter trials between the partner sites of the German Cancer Consortium and beyond.

### Knowledge Generated

The decentralized local execution of data analyses can solve many obstacles of cross-site collaboration. This work tackles organizational, legal, and technical challenges of distributed data analysis and shows its value in several use cases and studies.

### Relevance

Translating new developments into clinical practice is the ultimate goal of medical imaging research, and using these new technologies will yield enormous benefits for patients. The open-source Joint Imaging Platform presented in this work is realizing this step from the research laboratory into a real multicenter clinical study setting, supporting and enabling the translation of cutting-edge AI-based technologies into clinical practice.

sharing cumbersome but also necessary for generalizable image-based biomarkers and artificial intelligence (AI)–based image analysis.

Recent improvements in machine learning (ML) have enabled algorithms that can achieve results for specialized tasks equivalent to those of physicians and that are able to support human experts in improving their performance and efficiency.<sup>4,13-27</sup> These accomplishments elevated medical imaging to one of the most promising fields for practical application of ML in health care, aiming at better prediction, diagnosis, and treatment of diseases.<sup>13</sup> However, to enable a broad application of these techniques, a number of challenges still need to be overcome. The common denominator of all previous success stories is an extensive investment in collecting and curating a substantial amount of multicentric imaging data, which is critical to establish the required robustness of ML models. Thus, the obstacles for data sharing represent a bottleneck for medical research in general and for cancer research in particular.

To overcome this bottleneck, several projects and initiatives are working on facilitating, accelerating, and promoting collaboration in larger scientific networks. Existing platforms such as KETOS,<sup>28</sup> which is based on DataSHIELD<sup>29,30</sup> and targeted to perform statistical analysis on textual clinical information, were among the first to adopt the concept of federated on-site execution of algorithms to enable decentralized analysis of clinical data. The field of bioinformatics has brought up widely used platforms for standardized data analysis. Among these, Galaxy<sup>31</sup> is the most prominent solution for platform-based genome analyses. The Personal Health Train,<sup>32</sup> on the other hand, envisions federated scenarios that include the continuous federated training of ML models. In the area of medical imaging, Sharma et al<sup>33</sup> presented the Platform for Imaging in Precision Medicine (PRISM). PRISM focuses on the curation, management, and exploration of radiologic, pathologic, and clinical data collections, such as The

Cancer Imaging Archive,<sup>34</sup> not on the decentralized processing of imaging data and the realization of multicenter trials.

To fill this gap, the strategic initiative Joint Imaging Platform (JIP) was established by the German Cancer Consortium (DKTK).<sup>35</sup> The DKTK is a long-term initiative by the German Federal Ministry of Education and Research connecting more than 20 academic research institutes and university hospitals with the German Cancer Research Center to foster multicenter clinical trials for improved cancer diagnosis and treatment (Appendix Table A1). The consortium provides a unique opportunity for collecting large-scale high-quality imaging data from several institutions. The JIP is designed to facilitate collaborative imaging projects across institutions by addressing the typical technical, organizational, and legal challenges associated with the sharing of imaging data, acquisition parameters, analysis algorithms, or processing results. By enabling training, evaluation, and application of algorithms in large-scale federated clinical settings, the platform builds a solid and extensible foundation for federated learning scenarios. Leveraging open-source technologies, the JIP has the potential to serve as a promoter of prospective cross-center radiologic studies at unprecedented cohort sizes, not only within the DKTK but also beyond.

## METHODS

The JIP is designed as a federated data analysis and processing system (ie, for delivering methods and tools to the image data instead of collecting the data for processing and analysis). Strict on-site data processing mitigates common problems with data protection regulations because no personal data have to leave the clinic.

### Platform Requirements

We conducted a requirement analysis at each DKTK site to gather information regarding their specific demands and expectations. The collected responses revealed a quite

heterogeneous landscape considering patient count, modalities, and IT systems. Furthermore, the analysis revealed two main requirements for the JIP. First, enabling and supporting multicenter imaging studies was of utmost importance, including access to larger case numbers for retrospective data analyses and the facilitation of collaborations. Second, an improved integration of data processing, annotation, and sharing tools into the clinical environment was of interest, particularly ML and federated data analytics. These results were translated into the following individual aspects that should be realized in the platform.

**Integrability.** The fundamental principle of the JIP is based on the local execution of analysis methods as an extension of the existing clinical infrastructure. As a result, the platform should seamlessly integrate with existing local clinical systems, and the interaction of physicians with the JIP should be as close as possible to the established clinical workflows and tools.

**Data accessibility.** To achieve high compatibility with existing hospital systems, the widely established standard for Digital Imaging and Communications in Medicine (DICOM) should be used whenever possible. It should also be possible to view stored images and segmentations within the platform. Results of computations (eg, segmentations or parametric images) should be stored in DICOM to ensure high compatibility.

**Algorithmic accessibility.** The platform should facilitate the sharing and distribution of algorithmic developments between research groups across different sites. This requires a versatile and efficient integration path for in-house developments into the platform, supporting arbitrary processing steps using different programming languages and input and output formats.

**Data sovereignty.** Although the platform should enable joint projects within DKTK, it must put mechanisms into place that ensure full control over each site's local data.

**Data exploration.** Because algorithms are usually designed for specific types of input data, image properties such as modality, protocol, examined body parts, or patient characteristics are essential for selecting suitable training and test cases. The platform should enable users to filter appropriate data easily from existing image collections. The filters that identify a specific cohort should be shareable with partners.

**Scalability.** As a service-oriented application, the platform should ensure a high level of scalability. For example, an increasing demand should be responded by spawning more instances of a service that is under high load.

**Maintenance.** To ensure long-term sustainability of the platform, all platform instances in the consortium need to be kept up to date. Thus, to meet new developments and the continuous change in requirements, the platform must

be easy to maintain, update, and expand. The installation and maintenance of each individual instance must be possible by nonspecialized technical staff of each site.

### Platform Architecture

In response to these requirements, we designed a system that is structured into five building blocks (Fig 1). JIP SYSTEM realizes the technical basis and acts like an operating system of the platform. It takes care of provision, monitoring, and communication of services within the platform. JIP BASE consists of the more task-related components (eg, for the main user interface and authentication). JIP STORE contains components for data handling, management, and storage; JIP META contains components for metadata management, subject and image search, and selection; and JIP FLOW contains components for the controlled execution of processing sequences. Each of the functional units was realized based on open-source technologies. We have designed the system to be located within the protected hospital IT infrastructure. This allows for processing of the entire available patient data and facilitates integration into local procedures. The Data Supplement provides more detailed information about the individual components.

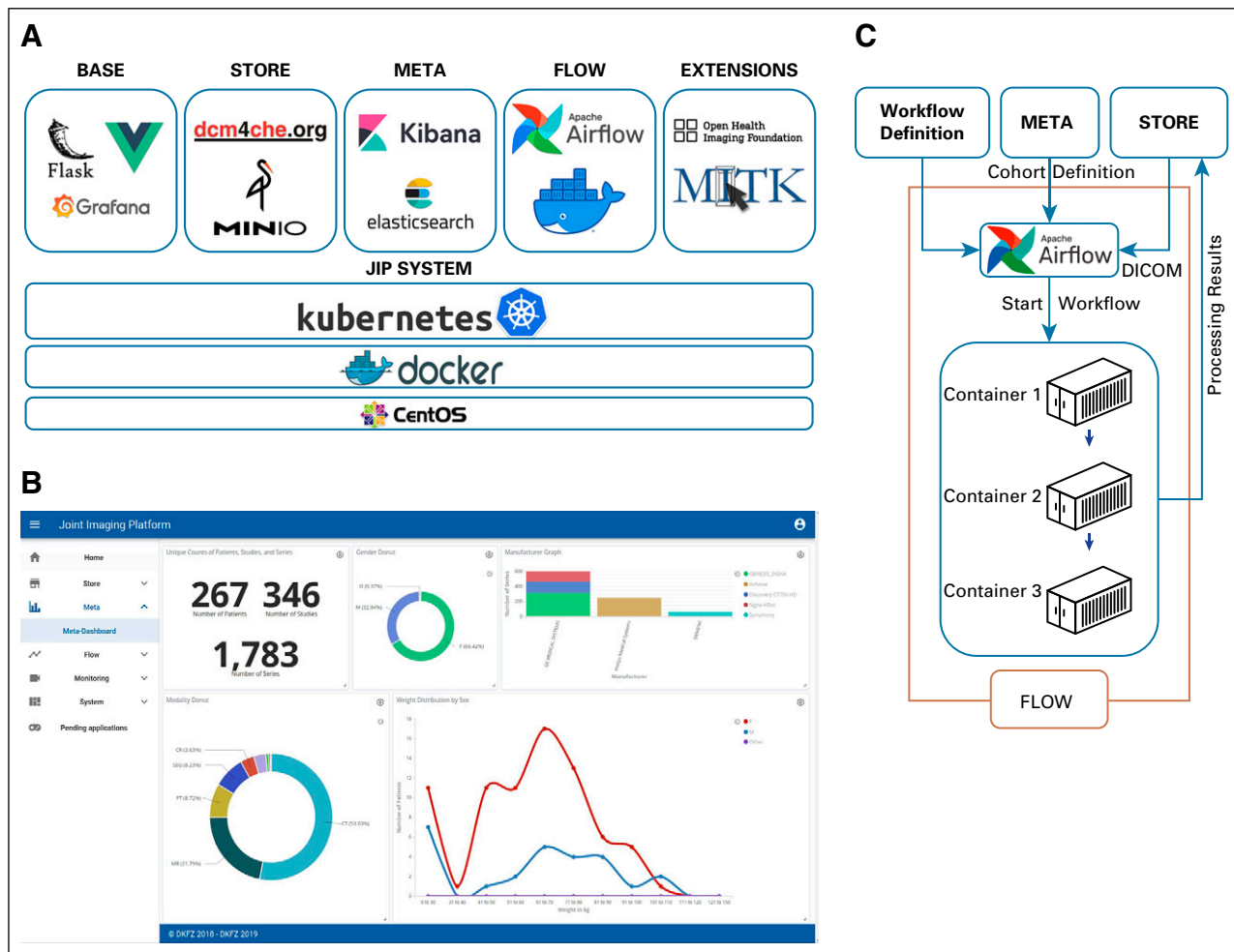
Because methods, technologies, and requirements in research are constantly evolving and the JIP is designed as an open platform for the community, extensibility of the platform for additional tools to explore, examine, or analyze medical data is crucial. Open interfaces, which are provided within the platform, allow a high degree of flexibility. This even extends to services that were originally not developed for use within a Web environment; the JIP offers Virtual Network Computing as an interface able to stream complete desktop applications to a Web browser. Some of the already integrated extensions (Fig 2D) are described in more detail in the Data Supplement. The seamless interaction of the JIP with other platform initiatives is also detailed in the Data Supplement.

## RESULTS

In this section, exemplary and complementary use cases that are realizable within the JIP and that cover all aspects of the previously defined platform requirements are described. The successful implementation of the JIP and its capabilities are further demonstrated in an overview of the current clinical and technical site involvements.

### Use Case 1: Data Harmonization and Integration

To enable comparability of study results, mutually agreeable scanner configurations or guidelines that achieve a more standardized imaging of certain patient groups are desirable in multicenter studies. The JIP supports such harmonization of data by enabling the development of algorithms that robustly handle multicenter data, including all the protocol and quality variations.



**FIG 1.** The Joint Imaging Platform (JIP) architecture and design principles. (A) Technology stack of the JIP: on the basis of any Linux-based operating system, the system components are realized as Docker containers and managed by Kubernetes. The actual application layer is structured in different functional units. (B) JIP BASE offers a common user interface, which is entirely Web based. The example shows the JIP META visualization of Digital Imaging and Communications in Medicine (DICOM) metadata in a Kibana dashboard. (C) Exemplary interplay between different JIP components: cohorts are defined in JIP META and handed over to JIP FLOW for data processing. Data input and output data are realized through communication with JIP STORE. (All trademarks and logos are the property of their respective owners.)

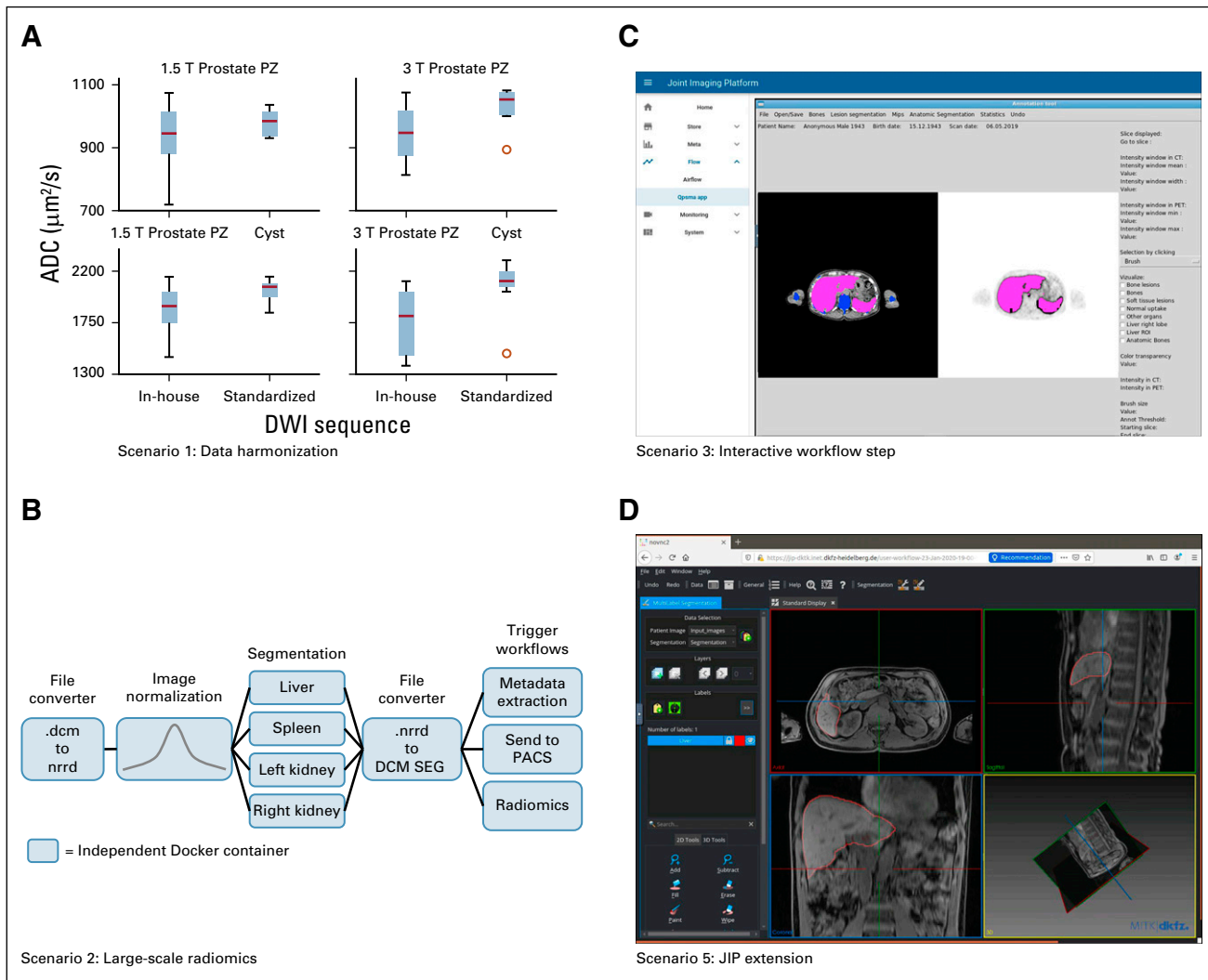
Within the DKTK consortium, the imaging protocols for diffusion-weighted magnetic resonance imaging of the brain and the prostate were standardized and validated using the JIP. The effect of the harmonization on the quantitative apparent diffusion coefficient measurements was demonstrated in comparison with canonical in-house sequences before and after standardization, exemplarily shown for prostate measurements (Fig 2A). Levene's test revealed that there were significant differences between the variances before and after harmonization at 3 T (peripheral zone [PZ],  $P = .004$ ; cyst,  $P = .00005$ ) but not at 1.5 T (PZ,  $P = .072$ ; cyst,  $P = .076$ ).

### Use Case 2: Automatic Large-Scale Radiomics Analysis

In this scenario, a large number of images are processed using fully automatic image quantification algorithms, more specifically a shape model-based organ segmentation

covering segmentations for kidneys, liver, and spleen in abdominal computed tomography (CT) scans<sup>36</sup> followed by a radiomics analysis of the resulting objects.<sup>37</sup>

Complex workflows are realized in the JIP by concatenating individual processing steps and pipelines. The generated segmentations are pushed into JIP STORE and trigger the subsequent radiomics workflow, which automatically extracts the radiomics features from the provided organ masks. All metadata of the generated DICOM-SEGs are automatically extracted and pushed into JIP META. The combination of Kubernetes and Apache Airflow allows automatic parallel execution of each pipeline instance across the configured computing cluster, reducing computation time through transparent parallelization. Figure 2B illustrates the fully automated processing workflow, which is formally defined as a directed acyclic graph with Docker containers as nodes.



**FIG 2.** Scenarios realized within the Joint Imaging Platform (JIP). (A) Comparative intraindividual magnetic resonance imaging measurements of a traveling volunteer across sites. Apparent diffusion coefficient (ADC) values in the prostate were analyzed in the peripheral zone (PZ) (top row) as well as in the prostate cyst (bottom row), using 1.5 T (left) and 3 T (right). Sequence standardization led to substantially reduced variance in comparison with in-house (not standardized) sequences. (B) Exemplary organ segmentation workflow as realized in scenario 2. (C) Interactive workflow component of the qPSMA software<sup>37a</sup> as realized in scenario 3. (D) Exemplary JIP extension offering a variety of tools for image segmentation. For this purpose, the Medical Imaging Interaction Toolkit (MITK) was wrapped in a Docker container and runs directly in the browser. DWI, diffusion-weighted imaging; PACS, picture archiving and communication system.

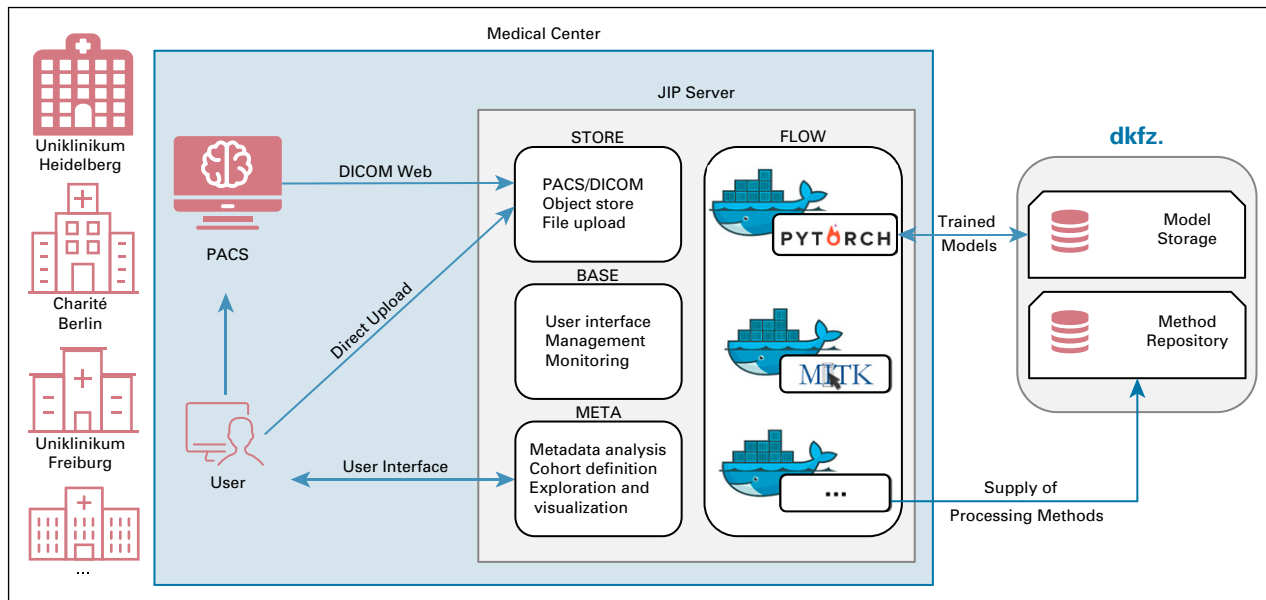
**Use Case 3: Interactive Analysis**

This scenario demonstrates that interactive desktop applications can be shipped via containers and integrated into otherwise automatic JIP processing workflows. The desktop application qPSMA<sup>37a</sup> for the interactive quantification of whole-body tumor volume of patients with prostate cancer using prostate-specific membrane antigen–positron emission tomography/CT images, developed by the Department of Nuclear Medicine at the Technical University Munich, was integrated into such a semiautomatic workflow. In this specific workflow, time-consuming preprocessing steps are automatically performed before the manual annotation step is triggered (Fig 2C). After the expert’s manual interaction, the automatic processing pipeline continues. No software

needs to be installed on the annotators’ workstations, and multiple instances can be started simultaneously, also within a workflow, allowing parallel and more complex annotation workflows that might involve multiple users.

**Use Case 4: Federated Data Analysis**

This use case focuses on the cross-site distribution and application of analysis tools (Fig 3). As shown in the previous use cases, once a tool is packaged in a container, it can be executed on the platform independently of its grade of automation or interactivity level, building the foundation for a federated data analysis. The Helmholtz incubator project Trustworthy Federated Data Analysis,<sup>38</sup> based on the JIP technology, is investigating trustworthy and



**FIG 3.** Clinical integration and federated setup of the Joint Imaging Platform (JIP). The JIP is located within the protected hospital information technology infrastructure, in principle allowing it to process the entire available patient data and facilitating its integration into local procedures. A central model storage was established at German Cancer Research Center to allow secure and access-controlled distribution and exchange of the platform itself, the analysis methods, and processing results in the form of containers. (All trademarks and logos are the property of their respective owners.) DICOM, Digital Imaging and Communications in Medicine; PACS, picture archiving and communication system.

regulatory compliant federated computing concepts. A hurdle in federated data analysis is the heterogeneity of data and metadata across sites, even when using standards such as DICOM. To approach this challenge, there is ongoing work to semiautomatically create mappings between different local conventions to a standardized metadata format. As a proof of concept, the development of this project will be validated in a federated radiation therapy study, aiming to reproduce results that were already generated and published by traditional means (ie, by pooling the data from all sites in the context of the DTKK Radiation Oncology Group study).<sup>39</sup>

#### Site User Engagement and Projects

Recently, developers at local sites have started to become involved and to migrate their processing workflows into the JIP. Measures to further encourage community involvement include a detailed developer guide and documentation,<sup>40</sup> together with an open-source release of the codebase.<sup>41</sup> The first JIP tech workshop has recently taken place, and further events such as hackathons and workshops are planned.

Since the initial release, several projects have started to or plan to use the platform. A subset of projects is listed in Table 1. The spectrum of applications reaches from simple data management to complex application of modern ML algorithms. For example, the surgical ARMANI trial will use the platform to investigate imaging associated with different resection strategies of liver metastases. In the course of the project outlined in case 3, the lesion load in prostate cancer will be examined.

#### DISCUSSION

The JIP provides a unified infrastructure across radiology and nuclear medicine departments of 10 university hospitals in Germany. The platform leverages state-of-the-art industry standards for cloud computing while adhering to on-premise hosting and execution. The technology stack is typical in modern cloud systems and can also easily be deployed using one of the leading commercial providers of cloud services.

The JIP offers standardized image processing by leveraging successfully established mechanisms from other fields. For example, we implemented a strictly browser-based interaction as in Galaxy<sup>31</sup> and decentralized execution of algorithms as it was suggested by KETOS for textual data,<sup>28</sup> while enabling the potential of container-based federated learning as suggested by the Personal Health Train.<sup>32</sup>

The decentralized approach enables compliance with data protection rules according to the European General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA). We provided support in the handling of data protection questions to the sites by distributing technical and organizational measures as required by the recently introduced GDPR. The decentralized data storage facilitates additional GDPR requirements such as consent management (including the handling of consent withdrawal), data transparency, and the right to rectification and erasure. Depending on the study protocol, data pseudonymization and anonymization can be designed directly into the processing workflow,

**TABLE 1.** Overview of Current and Prospective Multicenter Trials Using the JIP

Project Name	Description
ARMANI	A prospective trial to compare anatomic v wedge resection of liver metastases in patients with RAS-mutated colorectal cancer
MEMORI (ClinicalTrials.gov identifier: <a href="https://clinicaltrials.gov/ct2/show/study/NCT02287129">NCT02287129</a> )	Correlation of pre- and posttreatment PET/CT imaging biomarkers with histopathologic and molecular features in esophagogastric adenocarcinomas
NEOLAP (German Clinical Trials Register identifier: DRKS00019011)	Development and application of machine learning algorithms for noninvasive image-based subtyping of locally advanced pancreatic adenocarcinoma
PROACTIVE (German Clinical Trials Register identifier: DRKS00013915)	A prospective longitudinal examination of patients with low-risk prostate cancer aiming at investigating the benefits of multiparametric MRI for active surveillance
qPSMA-JIP	Retrospective evaluation of the prognostic value of PSMA-targeted PET/CT for clinical outcome in patients with metastatic hormone-sensitive and castration-resistant prostate cancer.
DKTK ROG <sup>39</sup>	Explorative study to predict locoregional control after postoperative chemoradiotherapy of locally advanced oropharyngeal carcinoma based on HPV-16 DNA status

Abbreviations: CT, computed tomography; DKTK, German Cancer Consortium; HPV, human papillomavirus; JIP, Joint Imaging Platform; MRI, magnetic resonance imaging; PET, positron emission tomography; PSMA, prostate-specific membrane antigen; ROG, Radiation Oncology Group.

facilitating the duties of the data controller as required by the GDPR. These principles naturally extend to the requirements of the HIPAA, such as central processing steps for personal health information management.

Because of regional differences, site-specific data protection concepts often represent specific challenges. For example, some clinical IT networks do not grant permanent external communication. This complicates scenarios like methods exchange or federated learning by adding manual steps and preventing automated updates. In addition, the server must be installed and maintained independently at each site. This significantly increases the need for reliability and automation of maintenance routines. We have addressed this issue on several levels, from the basic system architecture up to documentation and support measures. The first release cycles have shown that our measures were successful.

The integration of electronic health records (EHRs) will be evaluated in upcoming versions. The contained information might serve as an important pretest probability for certain medical conditions and should be available as input for ML algorithms. Additional nonimaging parameters can, for instance, be taken into account by linking to the DKTK CCP-IT.<sup>42,43</sup> Including radiologic reports from the radiologic information system presents an additional data source and allows for an even more powerful stratification of patient subgroups.

In the context of electronic case report forms and EHRs, the important aspect of data quality has been investigated extensively.<sup>44-47</sup> With a similar intention but with a focus on

imaging, future work on the JIP will include the development of automatic AI-based data quality assessment methods based on image metadata as well as the actual image data itself.

For a clinical application of AI-based methodologies, the trustworthiness of such techniques is also a key challenge. Trust can only be generated by a deep understanding of an algorithm's decision making, which can be promoted by new techniques of explainable AI, where the JIP as a research platform located in multiple university hospitals is the ideal tool to develop and test such approaches.<sup>27</sup>

To conclude, we have established a flexible decentralized analysis platform for medical images that respects data sovereignty and protects privacy by sharing algorithms instead of data. We observed that the availability of the JIP led to an unprecedented level of communication and collaboration within the radiologic and nuclear medicine research community of the DKTK. An increasing number of clinical studies have committed to use the JIP (Table 1). In addition, several requests for further extensions have been made (eg, supporting histopathology data).

Because DKTK is not the only collaborative network that is in need of a research imaging platform, the core implementation of the JIP will be available as an open-source software project named Kaapana.<sup>48</sup> By providing the platform and source code, we hope to mitigate the compatibility gap between systems in the heterogeneous clinical IT landscapes and lay the foundation for unprecedented research opportunities in data-driven medicine.

## AFFILIATIONS

<sup>1</sup>Division of Medical Image Computing, German Cancer Research Center, Heidelberg, Germany

<sup>2</sup>Medical Faculty Heidelberg, University of Heidelberg, Heidelberg, Germany

<sup>3</sup>German Cancer Consortium, Heidelberg, Germany

<sup>4</sup>Pattern Analysis and Learning Group, Radio-oncology and Clinical Radiotherapy, Heidelberg University Hospital, Heidelberg, Germany

<sup>5</sup>Division of Radiology, German Cancer Research Center, Heidelberg, Germany

<sup>6</sup>Klinik Diagnostische und Interventionelle Radiologie der Universität Heidelberg, Heidelberg, Germany

<sup>7</sup>Institut für Diagnostische und Interventionelle Radiologie, Universitätsklinikum Frankfurt, Frankfurt, Germany

<sup>8</sup>Klinik für Nuklearmedizin, Universitätsklinikum Frankfurt, Frankfurt, Germany

<sup>9</sup>Institut und Poliklinik für Diagnostische und Interventionelle Radiologie, Universitätsklinikum Carl Gustav Carus Dresden, Dresden, Germany

<sup>10</sup>Klinik und Poliklinik für Nuklearmedizin, Universitätsklinikum Carl Gustav Carus Dresden, Dresden, Germany

<sup>11</sup>Medical Faculty, Department of Diagnostic and Interventional Radiology, University Düsseldorf, Düsseldorf, Germany

<sup>12</sup>Klinik für Nuklearmedizin, Universitätsklinikum Düsseldorf, Düsseldorf, Germany

<sup>13</sup>Klinik für Diagnostische und Interventionelle Radiologie, Universitätsklinikum Tübingen, Tübingen, Germany

<sup>14</sup>Klinik für Nuklearmedizin und Klinische Molekulare Bildgebung, Universitätsklinikum Tübingen, Tübingen, Germany

<sup>15</sup>Department of Radiology, University Hospital, Ludwig Maximilian University Munich, Munich, Germany

<sup>16</sup>German Center of Lung Research, Giessen, Germany

<sup>17</sup>Klinik und Poliklinik für Nuklearmedizin, Klinikum der Universität München, München, Germany

<sup>18</sup>Institut für Diagnostische und Interventionelle Radiologie und Neuroradiologie, Universitätsklinikum Essen AÖR, Essen, Germany

<sup>19</sup>Klinik für Nuklearmedizin, Universitätsklinikum Essen AÖR, Essen, Germany

<sup>20</sup>Klinik für Radiologie (mit dem Bereich Kinderradiologie), Charité Universitätsmedizin Berlin, Berlin, Germany

<sup>21</sup>Klinik für Nuklearmedizin, Charité–Universitätsmedizin Berlin, Berlin, Germany

<sup>22</sup>Klinik und Poliklinik für Diagnostische und Interventionelle Radiologie, Universitätsmedizin Mainz, Mainz, Germany

<sup>23</sup>Klinik und Poliklinik für Nuklearmedizin, Universitätsmedizin Mainz, Mainz, Germany

<sup>24</sup>Institut für Diagnostische und Interventionelle Radiologie, Klinikum Rechts der Isar, Technical University of Munich, Munich, Germany

<sup>25</sup>Department of Computing, Imperial College London, London, United Kingdom

<sup>26</sup>Klinik und Poliklinik für Nuklearmedizin, Klinikum Rechts der Isar, Technical University of Munich, Munich, Germany

<sup>27</sup>Klinik für Diagnostische und Interventionelle Radiologie, Universitätsklinikum Freiburg, Freiburg, Germany

<sup>28</sup>Klinik für Nuklearmedizin, Universitätsklinikum Freiburg, Freiburg, Germany

<sup>29</sup>Klinische Kooperationseinheit Nuklearmedizin, Deutsches Krebsforschungszentrum Heidelberg, Heidelberg, Germany

<sup>30</sup>Universitätsmedizin Mannheim, Medizinische Fakultät Mannheim der Universität Heidelberg, Heidelberg, Germany

<sup>31</sup>Medizinische Physik in der Radiologie, Deutsches Krebsforschungszentrum Heidelberg, Heidelberg, Germany

## CORRESPONDING AUTHOR

Klaus Maier-Hein, PhD, German Cancer Research Center, Im Neuenheimer Feld 280, Heidelberg, 69120 Germany; e-mail: k.maier-hein@dkfz.de.

## EQUAL CONTRIBUTION

M.N., J.K., J.M., K.K., H-P.S., and K.M-H. contributed equally to this work.

## PRIOR PRESENTATION

Presented in part at the 101st Deutscher Röntgenkongress Presentation at the 1st German Cancer Research Congress, Heidelberg, Germany, February 4-5, 2019.

## SUPPORT

Supported by the German Cancer Consortium.

## AUTHOR CONTRIBUTIONS

**Conception and design:** Jonas Scherer, Marco Nolden, Jens Kleesiek, Jasmin Metzger, Klaus Kades, Verena Schneider, Oliver Sedlacek, Andreas M. Bucher, Thomas J. Vogl, Jens-Peter Kühn, Gerald Antoch, Konstantin Nikolaou, Wolfgang G. Kunz, Balthasar Schachtner, Jens Ricke, Alexander Radbruch, Lale Umutlu, Michael Forsting, Robert Seifert, Ken Herrmann, Philipp Mayer, Hans-Ulrich Kauczor, Winfried Brenner, Roman Kloeckner, Mathias Schreckenberger, Rickmer Braren, Georgios Kaissis, Marcus Makowski, Matthias Eiber, Andrei Gafita, Wolfgang A. Weber, Michael Bock, Fabian Bamberg, Jürgen Hennig, Tristan Kuder, Peter Neher, Ralf Floca, Heinz-Peter Schlemmer, Klaus Maier-Hein

**Administrative support:** Frank Grünwald, Jens-Peter Kühn, Christian la Fougère, Rickmer Braren, Jakob Neubauer, Heinz-Peter Schlemmer  
**Provision of study materials or patients:** Jens-Peter Kühn, Lars Schimmöller, Christian la Fougère, Peter Bartenstein, Bernd Hamm, Rickmer Braren, Jürgen Hennig, Uwe Haberkorn, Heinz-Peter Schlemmer

**Collection and assembly of data:** Marco Nolden, Jens Kleesiek, Klaus Kades, Michael Bach, Oliver Sedlacek, Andreas M. Bucher, Thomas J. Vogl, Frank Grünwald, Jens-Peter Kühn, Ralf-Thorsten Hoffmann, Oliver Bethge, Hans-Wilhelm Müller, Andreas Daul, Christian la Fougère, Wolfgang G. Kunz, Michael Ingris, Peter Bartenstein, Felix Nensa, Alexander Radbruch, Michael Forsting, Tobias Penzkofer, Bernd Hamm, Christoph Düber, Rickmer Braren, Rupert Trager, Jakob Neubauer, Jürgen Hennig, Philipp Tobias Meyer, Juri Ruf, Uwe Haberkorn, Klaus Maier-Hein

**Data analysis and interpretation:** Marco Nolden, Jens Kleesiek, Jasmin Metzger, Klaus Kades, Michael Bach, Andreas M. Bucher, Jens-Peter Kühn, Jörg Kotzerke, Lars Schimmöller, Christian la Fougère, Wolfgang G. Kunz, Alexander Radbruch, Lale Umutlu, Michael Forsting, Ken Herrmann, Tobias Penzkofer, Bernd Hamm, Roman Kloeckner, Rickmer Braren, Georgios Kaissis, Marco Reiser, Stefan O. Schoenberg, Peter Neher, Klaus Maier-Hein

**Manuscript writing:** All authors

**Final approval of manuscript:** All authors

**Accountable for all aspects of the work:** All authors

## AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to [www.asco.org/rwc](http://www.asco.org/rwc) or [ascopubs.org/cci/author-center](http://ascopubs.org/cci/author-center).



Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians ([Open Payments](#)).

**Andreas M. Bucher**

**Honoraria:** Bayer, Guebert

**Travel, Accommodations, Expenses:** Bayer, Guebert

**Frank Grünwald**

**Honoraria:** Henning

**Konstantin Nikolaou**

**Honoraria:** Siemens Healthineers, Bayer Schering Pharma

**Consulting or Advisory Role:** Siemens Healthineers

**Speakers' Bureau:** Siemens Healthineers, Siemens Healthineers (Inst)

**Research Funding:** Bayer Schering Pharma (Inst)

**Travel, Accommodations, Expenses:** Siemens Healthineers, Bayer Schering Pharma

**Michael Ingrisich**

**Stock and Other Ownership Interests:** Siemens Healthineers

**Jens Ricke**

**Honoraria:** Sirtex Medical

**Consulting or Advisory Role:** Ipsen

**Research Funding:** Sirtex Medical (Inst)

**Travel, Accommodations, Expenses:** BTG

**Alexander Radbruch**

**Honoraria:** Bayer (Inst), Guerbet (Inst), Bayer, Guerbet, Novartis

**Consulting or Advisory Role:** Guerbet, Bayer

**Speakers' Bureau:** Bayer, Guerbet

**Research Funding:** Bayer (Inst), Guerbet (Inst)

**Travel, Accommodations, Expenses:** Bayer, Guerbet

**Lale Umutlu**

**Consulting or Advisory Role:** Siemens Healthineers

**Speakers' Bureau:** Siemens Healthineers, Bayer

**Research Funding:** Siemens Healthineers (Inst)

**Travel, Accommodations, Expenses:** Bayer, Siemens Healthineers

**Other Relationship:** Vara

**Uncompensated Relationships:** Vara

**Ken Herrmann**

**Leadership:** Sofie Biosciences

**Stock and Other Ownership Interests:** Sofie Biosciences

**Consulting or Advisory Role:** Novartis, Bain Capital, Bayer, Adacap, Amgen, BTG, Ipsen, ITG, ROTOP (Inst), Siemens Healthineer, GE Healthcare

**Hans-Ulrich Kauczor**

**Speakers' Bureau:** Philips Healthcare, Boehringer Ingelheim, MSD, AstraZeneca

**Research Funding:** Siemens Healthineers (Inst), Philips Healthcare (Inst), Bayer Schering Pharma (Inst)

**Tobias Penzkofer**

**Research Funding:** AGO (Inst), Aprea AB (Inst), ARCAGY-GINECO (Inst), Astellas Pharma Global (APGD) (Inst), AstraZeneca (Inst), Clovis Oncology (Inst), Dohme (Inst), Holaira (Inst), Incyte (Inst), Karyopharm Therapeutics (Inst), Lion Biotechnologies (Inst), MedImmune (Inst), Merck Sharp (Inst), Millennium (Inst), Morphotec (Inst), NovoCure (Inst), PharmaMar S.A. and PharmaMar USA (Inst), Roche (Inst), Siemens Healthineers (Inst), Tesaro (Inst)

**Bernd Hamm**

**Honoraria:** Bayer Schering Pharma

**Consulting or Advisory Role:** Canon, InnoRa GMBH

**Research Funding:** Abbot/AbbVie (I), Bracco Diagnostics (I), Guerbet (I), Novartis (I), GE Healthcare (I), Siemens (I)

**Other Relationship:** Board member of professional societies

**Roman Kloeckner**

**Consulting or Advisory Role:** Boston Scientific, Bristol Myers Squibb, Guerbet, Roche

**Speakers' Bureau:** BTG, Bristol Myers Squibb, Guerbet, Siemens Healthineers

**Mathias Schreckenberger**

**Honoraria:** Takeda

**Matthias Eiber**

**Consulting or Advisory Role:** Blue Earth Diagnostics, ABX Advanced Biochemical Compounds

**Research Funding:** Siemens, Blue Earth Diagnostics, ABX Advanced Biochemical Compounds

**Patents, Royalties, Other Intellectual Property:** Patent application for rhPSMA

**Travel, Accommodations, Expenses:** Bayer Schering Pharma

**Wolfgang A. Weber**

**Consulting or Advisory Role:** ITG, Blue Earth Diagnostics (Inst), Pentixapharm (Inst)

**Research Funding:** Blue Earth Diagnostics (Inst), ITG (Inst)

**Patents, Royalties, Other Intellectual Property:** Patent for PARP imaging agent

**Travel, Accommodations, Expenses:** Blue Earth Diagnostics, Pentixapharm

**Michael Bock**

**Research Funding:** Siemens Healthineers

**Fabian Bamberg**

**Speakers' Bureau:** Siemens Healthineers, Bracco Diagnostics, Bayer Health

**Research Funding:** Siemens Healthineers (Inst), Bayer Health (Inst)

**Philipp Tobias Meyer**

**Consulting or Advisory Role:** GE Healthcare, OPASCA

**Juri Ruf**

**Employment:** Gilead Sciences (I)

**Uwe Haberkorn**

**Patents, Royalties, Other Intellectual Property:** Patent for FAP inhibitors for nuclear medicine imaging and therapy (Inst)

**Stefan O. Schoenberg**

**Other Relationship:** Siemens Healthineers (Inst)

**Tristan Kuder**

**Patents, Royalties, Other Intellectual Property:** Patents for magnetic resonance diffusion-weighted imaging phantoms.

**Uncompensated Relationships:** HQ Imaging GmbH

**Heinz-Peter Schlemmer**

**Honoraria:** Bayer/Vital, Siemens Healthineers, Bracco Diagnostics, Curagita

**Consulting or Advisory Role:** Siemens Healthineers, Bracco Diagnostics

**Research Funding:** Siemens Healthineers (Inst), Profound (Inst)

**Travel, Accommodations, Expenses:** Siemens Healthineers, Bayer/Vital, Curagita, Bracco Diagnostics

**Klaus Maier-Hein**

**Research Funding:** Siemens Healthineers

No other potential conflicts of interest were reported.

## REFERENCES

1. Savadjiev P, Chong J, Dohan A, et al: Image-based biomarkers for solid tumor quantification. *Eur Radiol* 29:5431-5440, 2019
2. Parmar C, Grossmann P, Bussink J, et al: Machine learning methods for quantitative radiomic biomarkers. *Sci Rep* 5:13087, 2015
3. Kurtz DM, Esfahani MS, Scherer F, et al: Dynamic risk profiling using serial tumor biomarkers for personalized outcome prediction. *Cell* 178:699-713.e19, 2019
4. Kickingereder P, Isensee F, Tursunova I, et al: Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks: A multicentre, retrospective study. *Lancet Oncol* 20:728-740, 2019
5. Aerts HJWL: The potential of radiomic-based phenotyping in precision medicine: A review. *JAMA Oncol* 2:1636-1642, 2016
6. Amin S, Bathe OF: Response biomarkers: Re-envisioning the approach to tailoring drug therapy for cancer. *BMC Cancer* 16:850, 2016
7. Harry VN, Semple SI, Parkin DE, et al: Use of new imaging techniques to predict tumour response to therapy. *Lancet Oncol* 11:92-102, 2010
8. Gillies RJ, Kinahan PE, Hricak H: Radiomics: Images are more than pictures, they are data. *Radiology* 278:563-577, 2016
9. Rocher L, Hendrickx JM, de Montjoye Y-A: Estimating the success of re-identifications in incomplete datasets using generative models. *Nat Commun* 10:3069, 2019
10. Bartling S, Friesike S (eds): Challenges of open data in medical research, in *Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*. Cham, Switzerland, Springer International Publishing, 2014, pp 297-307
11. Ravindra V, Grama A: De-anonymization attacks on neuroimaging datasets. <http://arxiv.org/abs/1908.03260>
12. van Panhuis WG, Paul P, Emerson C, et al: A systematic review of barriers to data sharing in public health. *BMC Public Health* 14:1144, 2014
13. Topol EJ: High-performance medicine: The convergence of human and artificial intelligence. *Nat Med* 25:44-56, 2019
14. Rajpurkar P, Irvin J, Zhu K, et al: CheXNet: Radiologist-level pneumonia detection on chest x-rays with deep learning. <http://arxiv.org/abs/1711.05225>
15. Gale W, Oakden-Rayner L, Carneiro G, et al: Detecting hip fractures with radiologist-level performance using deep neural networks. <http://arxiv.org/abs/1711.06504>
16. Schelb P, Kohl S, Radtke JP, et al: Classification of cancer at prostate MRI: Deep learning versus clinical PI-RADS assessment. *Radiology* 293:607-617, 2019
17. Esteva A, Kuprel B, Novoa RA, et al: Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542:115-118, 2017
18. Kuo W, Häne C, Mukherjee P, et al: Expert-level detection of acute intracranial hemorrhage on head computed tomography using deep learning. *Proc Natl Acad Sci USA* 116:22737-22745, 2019
19. Hekler A, Utikal JS, Enk AH, et al: Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. *Eur J Cancer* 118:91-96, 2019
20. Campanella G, Hanna MG, Geneslaw L, et al: Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med* 25:1301-1309, 2019
21. Ardila D, Kiraly AP, Bharadwaj S, et al: End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med* 25:954-961, 2019
22. Christiansen EM, Yang SJ, Ando DM, et al: In silico labeling: Predicting fluorescent labels in unlabeled images. *Cell* 173:792-803.e19, 2018
23. Tomašev N, Glorot X, Rae JW, et al: A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* 572:116-119, 2019
24. Kleesiek J, Morshuis JN, Isensee F, et al: Can virtual contrast enhancement in brain MRI replace gadolinium? A feasibility study. *Invest Radiol* 54:653-660, 2019
25. Poplin R, Varadarajan AV, Blumer K, et al: Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng* 2:158-164, 2018
26. Lu MT, Ivanov A, Mayrhofer T, et al: Deep learning to assess long-term mortality from chest radiographs. *JAMA Netw Open* 2:e197416, 2019
27. Towards trustable machine learning. *Nat Biomed Eng* 2:709-710, 2018
28. Gruendner J, Schwachhofer T, Sippl P, et al: KETOS: Clinical decision support and machine learning as a service—A training and deployment platform based on Docker, OMOP-CDM, and FHIR Web Services. *PLoS One* 14:e0223010, 2019
29. Gaye A, Marcon Y, Isaeva J, et al: DataSHIELD: Taking the analysis to the data, not the data to the analysis. *Int J Epidemiol* 43:1929-1944, 2014
30. Wilson RC, Butters OW, Avraam D, et al: DataSHIELD: New directions and dimensions. *Data Sci J* 16:21, 2017
31. Goecks J, Nekutenko A, Taylor J: Galaxy: A comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* 11:R86, 2010
32. van Soest J, Sun C, Musmann O, et al: Using the personal health train for automated and privacy-preserving analytics on vertically partitioned data. *Stud Health Technol Inform* 247:581-585, 2018
33. Sharma A, Tarbox L, Kurc T, et al: PRISM: A platform for imaging in precision medicine. *JCO Clin Cancer Inform* 4:491-499, 2020
34. Clark K, Vendt B, Smith K, et al: The Cancer Imaging Archive (TCIA): Maintaining and operating a public information repository. *J Digit Imaging* 26:1045-1057, 2013
35. Joos S, Nettelbeck DM, Reil-Held A, et al: German Cancer Consortium (DKTK): A national consortium for translational cancer research. *Mol Oncol* 13:535-542, 2019
36. Norajitra T, Maier-Hein KH: 3D statistical shape models incorporating landmark-wise random regression forests for omni-directional landmark detection. *IEEE Trans Med Imaging* 36:155-168, 2017
37. Götz M, Nolden M, Maier-Hein K: MITK phenotyping: An open-source toolchain for image-based personalized medicine with radiomics. *Radiother Oncol* 131:108-111, 2019
- 37a. Gafita A, Bieth M, Krönke M, et al: qPSMA: Semiautomatic software for whole-body tumor burden assessment in prostate cancer using 68Ga-PSMA11 PET/CT. *J Nucl Med Off Publ Soc Nucl Med* 60(9):1277-1283, 2019
38. Helmholtz Center for Information Security, German Cancer Research Center: Trustworthy Federated Data Analytics (TFDA). <https://tfda.hmsp.center>
39. Lohaus F, Linge A, Tinhofer I, et al: HPV16 DNA status is a strong prognosticator of loco-regional control after postoperative radiochemotherapy of locally advanced oropharyngeal carcinoma: Results from a multicentre explorative study of the German Cancer Consortium Radiation Oncology Group (DKTK-ROG). *Radiother Oncol* 113:317-323, 2014
40. German Cancer Consortium: Joint Imaging Platform. <https://jip.dtk.dkfz.de/jiphomepage/>
41. Github: Kaapana repository. <https://github.com/kaapana/kaapana>
42. Lablans M: Konzept der CCP-IT des DKTK. [https://dtk.dkfz.de/application/files/1814/6235/8457/Konzept\\_CCP-IT.pdf](https://dtk.dkfz.de/application/files/1814/6235/8457/Konzept_CCP-IT.pdf)

43. Lablans M, Schmidt EE, Ückert F: An architecture for translational cancer research as exemplified by the German Cancer Consortium. *JCO Clin Cancer Inform* 2:1-8, 2018
44. Zaccaria GM, Ferrero S, Rosati S, et al: Applying data warehousing to a phase III clinical trial from the Fondazione Italiana Linfomi Ensures superior data quality and improved assessment of clinical outcomes. *JCO Clin Cancer Inform* 3:1-15, 2019
45. Weiskopf NG, Bakken S, Hripcsak G, et al: A data quality assessment guideline for electronic health record data reuse. *EGEMS (Wash DC)* 5:14, 2017
46. Terry AL, Stewart M, Cejic S, et al: A basic model for assessing primary health care electronic medical record data quality. *BMC Med Inform Decis Mak* 19:30, 2019
47. Kahn MG, Callahan TJ, Barnard J, et al: A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data. *EGEMS (Wash DC)* 4:1244, 2016
48. Kaapana. <https://kaapana.ai/>



## APPENDIX

**TABLE A1.** Key Facts About the DKTK

The discovery and quick development of new, innovative, and effective treatments is our response to increasing costs and increasing demand for cancer care. For this reason, a new German Health Research Center, the DKTK, was established in 2012, with > 20 academic research institutes and university hospitals at eight sites and with the German Cancer Research Center (DKFZ) as the consortium's core center. The DKTK is a long-term, joint initiative involving the German Federal Ministry of Education and Research, participating German states, and the DKFZ. It aims to discover, develop, test, and apply new personalized oncology strategies, with a focus on bringing new, promising results from research into clinical development and application as fast as possible.

## Key facts about the DKTK

> 950 medical and scientific researchers in clinically oriented cancer research

Installation of a high-quality and long-term research infrastructure across all partner sites

Focus on multicenter projects, complying with strict data protection standards

Abbreviations: DKFZ, German Cancer Research Center; DKTK, German Cancer Consortium.