

Features, Data, and Algorithms

# AI-Rad Companion Chest CT VA13

Whitepaper – September 2021



## Table of contents

Introduction .....	3
Product features .....	3
Workflow .....	6
Algorithm description .....	7
Data requirements .....	10
Proof points: performance and clinical value .....	12
References .....	15

## Introduction

AI-Rad Companion Chest CT is a decision support tool for the radiological assessment of computed tomography (CT) images of the thorax. It helps radiologists interpret CT images of the thorax more quickly and more precisely (doing more by doing less) and reduces the time needed to document the findings with the help of automatic measurements. It is vendor-neutral, which means that the software can evaluate image data from any CT system manufacturer. Enabled by the teamplay digital health platform and using state-of-the-art image processing algorithms supported by artificial intelligence, AI-Rad Companion Chest CT delivers value in four key areas:

1. Accelerated interpretation and workflow efficiency
2. Improved clinical outcomes and increased accuracy
3. Provision of additional clinically relevant information and visual highlighting
4. Standardized results while helping to reduce inter-reader variability

It focuses on three main parts of the thorax: the lungs (AI-Rad Companion (Pulmonary)), the cardiovascular system (AI-Rad Companion (Cardiovascular)) and the spine (AI-Rad Companion (Musculoskeletal)).

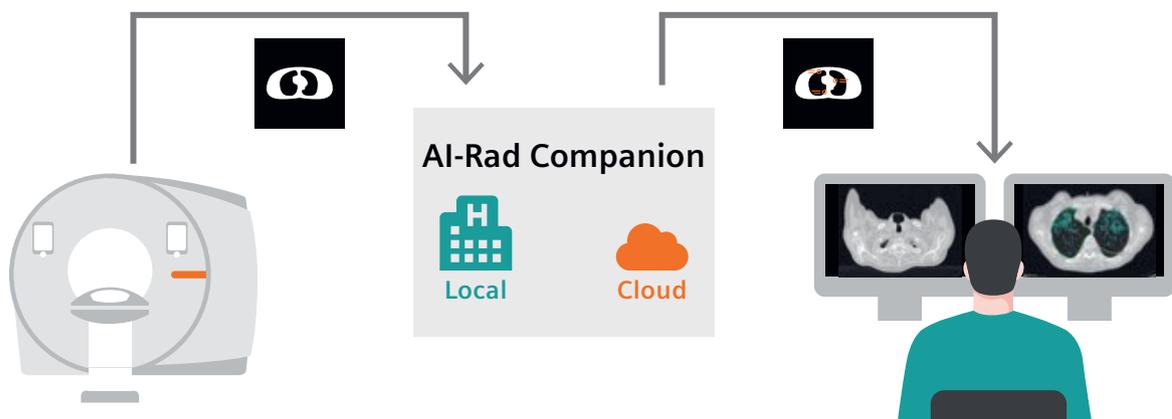
The typical workflow consists out of the following steps:

1. Reconstructed CT images of the thorax are sent to the PACS for interpretation.
2. In parallel, they are also sent to AI-Rad Companion. Extensions are launched automatically.
3. The results of AI-Rad Companion can either be sent to a web-based review software or directly to PACS. Here they can be used in combination with the original data for reporting purposes.

This whitepaper is intended to provide an overview of the product features, describe the individual algorithmic components, discuss requirements for data to be processed using the device, and to provide an overview of internal and external proof points assessing the performance and illustrating the clinical value of the application.

## Product features

AI-Rad Companion Chest CT consists of three medical devices: AI-Rad Companion (Pulmonary), AI-Rad Companion (Cardiovascular) and AI-Rad Companion (Musculoskeletal).



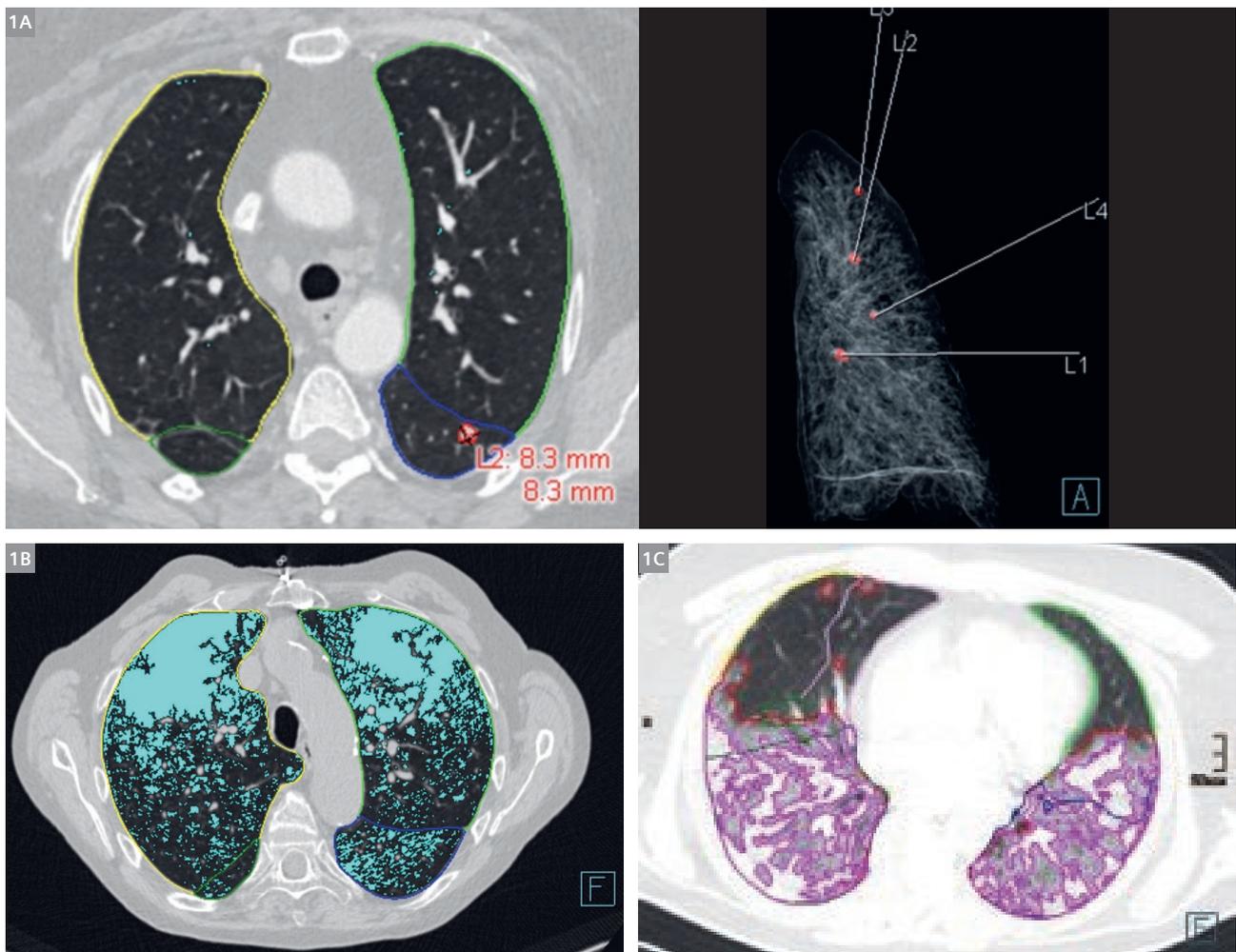
## AI-Rad Companion (Pulmonary)

The radiological assessment of pulmonary nodules is one of the most common indications for Chest CT. The radiologist needs to identify the nodules, measure the diameters and – ideally – also their volume (see guidelines of the Fleischner Society [1]). A second important application of Chest CT is the analysis of lung parenchyma. Reduced density can indicate emphysema and/or COPD while increased density can indicate inflammatory processes such as pneumonia.

AI-Rad Companion (Pulmonary) provides the following features with respect to the analysis of the lung:

- Detection and segmentation of lung nodules and localization with respect to lung lobes
- Analysis of the lung parenchyma based on segmented lung lobes with respect to:
  - areas of low attenuation (low attenuation volume, or LAV)
  - areas of opacity
  - volume of lung lobes

Exemplary outputs of AI-Rad Companion (Pulmonary) are shown in Figure 1. In the product, the LAV-Analysis is called “Lung Parenchyma Analysis”, while the opacity analysis is called “Pulmonary Density”<sup>1</sup>.



**Figure 1:** Outputs of the Pulmonary feature. Lung nodule detection and measurement (1A), LAV-analysis (1B), opacity detection (1C).

<sup>1</sup> The Pulmonary Density feature is new in VA12A without FDA clearance. According to FDA policy “Enforcement Policy for Imaging Systems During the Coronavirus Disease 2019 (COVID-19) Public Health Emergency” issued in April 2020, the manufacturer is allowed to market this feature without FDA clearance. This policy is intended to remain in effect only for the duration of the public health emergency related to COVID-19 declared by the HHS, including any renewals made by the HHS Secretary in accordance with section 319(a)(2) of the Public Health Services Act (42 U.S.C. 247d(a)(2)). Pulmonary Density results are not indicated for the diagnosis of COVID-19. Only in vitro diagnostic testing is currently the definitive method to diagnose COVID-19.

## AI-Rad Companion (Cardiovascular)

For the radiological assessment of the cardiovascular system a large variety of dedicated CT scan protocols exist depending on the clinical indication. The protocols differ mainly with respect to the cardiac phase in which the acquisition is performed (controlled via ECG-gating) and the type and timing of contrast enhancement. AI-Rad Companion is designed to work with any of these protocols, particularly the most generic non-gated and non-contrast-enhanced Chest CT scans. Of course, this also limits the analysis to features that can be reliably assessed on generic chest CT data. The features are:

- Measurement of heart volume and coronary calcium volume (on unenhanced data only)
- Segmentation of aorta and diameter measurements at 9 landmarks according to AHA-guidelines [2] (on both native and contrast-enhanced data)

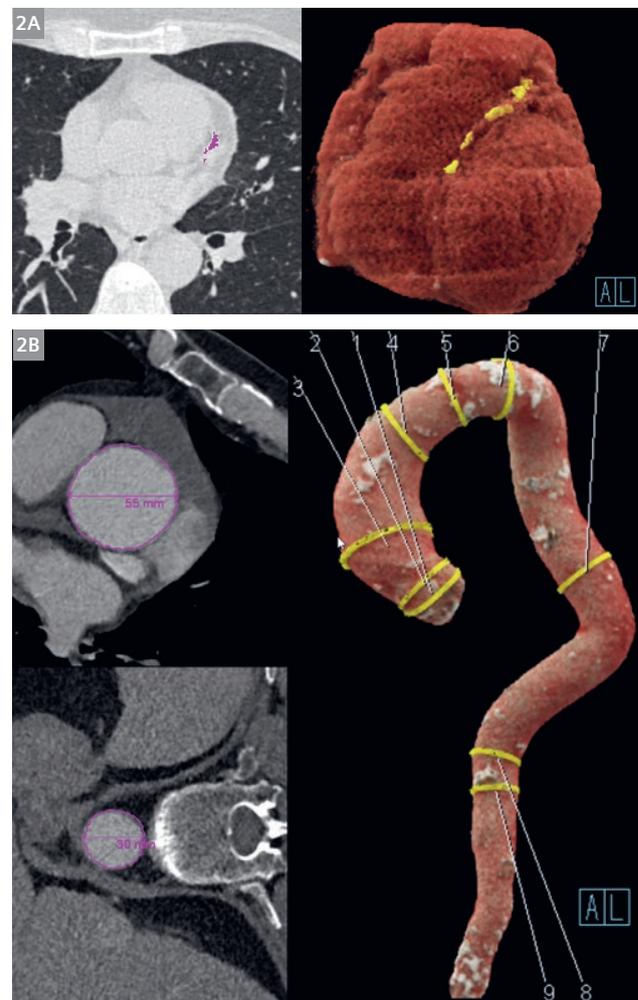
Exemplary outputs of AI-Rad Companion (Cardiovascular) are shown in Figure 2.

It is important to understand that the user should interpret the results of AI-Rad Companion (Cardiovascular) with respect to the actual scan protocol used. E.g., motion artifacts on a non-gated exam may hamper the accuracy of the aortic diameter measurements. Analogously, the coronary calcium analysis provides the total volume of the – potentially motion corrupted – calcium clusters but does not perform Agatston scoring which requires a gated scan and is the gold standard for dedicated cardiac CT scans.

However, the importance of the analysis of both coronary calcium and aorta in the context of chest CT is to be pointed out. Both features are listed in the recommendations by the ACR Incidental Findings Committee [3]. The 2016 SCCT/STR guidelines [4] state that coronary artery calcium “should be evaluated and reported on all non-contrast chest CT examinations”. Analogously, in a consensus statement the British societies BSCI/BSCCT and BSTI [5] “recommend that coronary artery calcification is reported on all non-gated thoracic CT using a simple patient-based score (none, mild, moderate, severe)”.

In their 2010 guidelines [2] the ACCF/AHA states that “many thoracic aortic diseases, results of treatment for stable, often asymptomatic, but high-risk conditions

are far better than the results of treatment required for acute and often catastrophic disease presentations. Thus, the identification and treatment of patients at risk for acute and catastrophic disease presentations prior to such an occurrence are paramount to eliminating the high morbidity and mortality associated with acute presentations” and hence motivates the automatic analysis of the thoracic aorta on any chest CT. In the guidelines it is also described that the thoracic aorta should be measured at nine predefined anatomical locations. “The use of standardized measurements helps minimize errant reports of significant aneurysm growth due to technique or interreader variability in measuring technique.” [2]



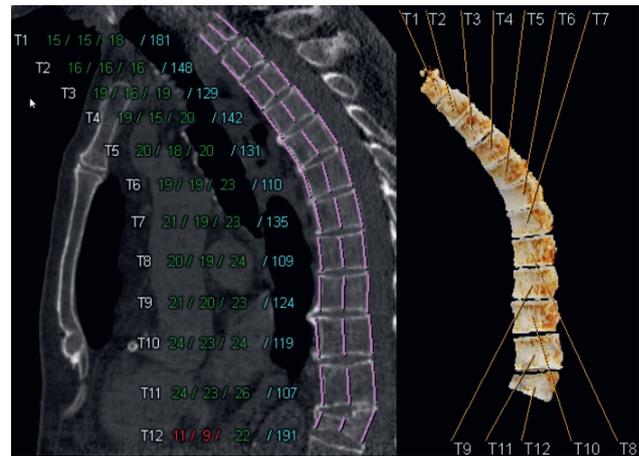
**Figure 2:** Outputs of the Cardiovascular device. Coronary calcium detection (2A), aorta analysis (2B).

## AI-Rad Companion (Musculoskeletal)

Osteoporosis manifests as loss of bone density e.g. in the spine, and consequently in vertebral compression fractures. The International Osteoporosis Foundation (IOF) states that “there is strong evidence of widespread under-diagnosis of vertebral fractures” [6]. Pickhardt et al. [7] and more recently Cohen et al. [8] showed that the HU-values of the spine obtained from CT data acquired for other indications can be used to identify osteoporotic patients and called this approach “opportunistic screening for osteoporosis”. AI-Rad Companion (Musculoskeletal) provides:

- Labeling and segmentation of thoracic vertebrae
- Measurements of vertebrae heights
- Mean vertebral body HU-density

Exemplary outputs are shown in Figure 3.



**Figure 3:** Output of the Musculoskeletal device: Height and density measurements of the thoracic vertebrae.

## Workflow

AI-Rad Companion Chest CT offers advanced ways of workflow customization. By design, all results are presented in the form of an annotated axial series, a 3D rendering, and a concise summary table – enabling seamless integration into a variety of different reading workflows. Moreover, a DICOM Structured Report with measurement results is provided.

1. Efficiency gains are best accomplished when AI-Rad Companion is used to automate the repetitive and manually tedious task such as measurements. All results are readily available the moment a case is

opened. A results table summarizes all findings and measurements. A color-coding scheme is used to draw the attention to potential abnormalities. Added 3D renderings quickly provide an intuitive presentation overview of the type, number, and spatial context of all findings. Upon confirmation of the findings, the results are straightforwardly transferred to the report.

2. Results of the AI-Rad Companion are best incorporated into the primary read by synchronizing annotated axials with the corresponding original series. As the reader scrolls through the stack, through highlighting



findings their attention is drawn to potential abnormalities. At the same time, the correctness of the AI results is easily verified through comparison with unannotated series.

3. Minimal disruption to the established workflow and unbiased reading is achieved when AI is used in a “spell checker” mode. Here, the reader would stick to their established reading patterns, but just before closing the case one last glance at the results pictogram allows for a quick and easy confirmation that indeed nothing was missed.



## Algorithm description

### Lung nodule detection (lung CAD) and segmentation

Lung CAD processing is performed in several consecutive steps: Preprocessing, Candidate Generation, Classification, and Post-Filtering.

In the Preprocessing step the input image is standardized, and parenchyma is segmented using specialized Convolutional Neural Net (CNN, V-Net). This allows restricting the detection of findings within the lungs while optimizing the computation time.

Candidate Generation aims at achieving high sensitivity while keeping the number of candidates to a manageable number. The isotropic volume is partitioned into sub-volumes that are processed using a CNN. Then, filtering and non-maximum suppression yield a list of candidates for each sub-volume. A predefined threshold is applied on the confidence score to remove the least confidence (low score) findings. All candidates exceeding this threshold are passed on to the Classification step.

Classification utilizes a CNN-based classifier to process each candidate. The classifier calculates the feature values for each candidate and uses a soft-max function to estimate the likelihood of its type as either “nodule” or “non-nodule.” Candidates meeting or exceeding

a likelihood value (a final confidence score above or equal to an empirically determined threshold) are labeled as nodule candidates.

The Post-Filtering step includes the application of two cascaded filters. The first one aims at removing false positives originating **(a) from the colon** and a second one from **(b) calcified protrusions** (for example, areas where the sternum meets the manubrium, spine malformations, and osteophytes, and so on).

The first filter is a CNN-based classifier that has a similar structure to that of the classifier in the Classification step. The second filter uses three orthogonal slices at the candidate location as input to three CNN-based classifiers (one per slice). The results from the three classifiers is then combined by a max-voting mechanism. Any candidate deemed a false positive by either filter is thus removed.

The algorithms have been trained using more than 2,000 manually curated CT data sets. Network layout diagrams have been published by Chamberlin et al. [9].

After detection, nodules are segmented by an algorithm based on region growing. Diameter and volume measurements are provided.

## Lung lobe segmentation

The lung lobe segmentation algorithm computes segmentation masks of the five lung lobes for a given CT data set of the chest. First, multi-scale deep reinforcement learning [10] is used to robustly detect anatomical landmarks in a CT volume. The carina bifurcation and/or sternum tip are used to identify the lung region of interest (ROI). Next, the lung ROI image is resampled to a 2 mm isotropic volume and fed into an adversarial Deep Image-to-Image Network (DI2IN) [11] to generate the lung segmentation. Finally, the ROI segmentation mask is remapped to have the same dimension and the resolution as the input data. The DI2IN has been first trained on over 8,000 CT scans from a large group of patients with various diseases, then fine-tuned with over 1,000 scans with abnormal patterns including interstitial lung disease (ILD), pneumonia, and COVID-19.

The volume of the individual lung lobes, the left and right lung and of the complete lung are reported.

## LAV analysis

The LAV analysis is threshold-based, i.e. the algorithm determines all voxels below -950 **Hounsfield Units (HU)** in the lung. The threshold of -950 HU is widely used for the quantification of emphysema [12]. For each lung lobe as well as for the complete lung (i.e. the combination of all lobes) the ratio of LAV (LAV %) is reported. The following thresholds are being used as default values for highlighting:

Category	Condition
I	$LAV \% < 12.5 \%$
II	$12.5 \% \leq LAV \% < 25 \%$
III	$25 \% \leq LAV \% < 37.5 \%$
IV	$LAV \% \geq 37.5 \%$

**Table 1:** Thresholds for LAV %.

## Opacity detection and quantification

The detection and quantification of opaque regions in the lung – typically associated to viral pneumonia such as Covid-19 – uses a DenseUNet [13] with anisotropic kernels. Details of the algorithm are described by Chaganti et al. [14]. The algorithm has been trained on over 900 CT scans from patients with ILDs, pneumonia, and COVID-19.

The detected opacities are quantified by computing the percentage of opacity (PO, per lobe and per lung) and the percentage of high opacities (PHO, by applying a threshold of -200 HU on the subset of opaque regions). Based on the PO a lung severity score (LSS) is calculated according to Bernheim et al. [15]:

LSS	Condition
0	$PO = 0$
1	$0 < PO \leq 25 \%$
2	$25 \% < PO \leq 50 \%$
3	$50 \% < PO \leq 75 \%$
4	$PO > 75 \%$

**Table 2:** Thresholds for PO.

A total LSS is computed as the sum of the individual scores per lobe.

## Heart segmentation

The heart segmentation is performed using a deep U-shaped network [13] consisting of four convolutions and down-sampling steps, followed by four similar up-sampling layers. It has been trained on over 650 CT data sets. Subsequently, the heart segmentation mask is used to compute the heart volume.

## Coronary calcium detection

Using the heart mask as ROI an initial set of voxels as candidates for potentially calcified regions is obtained by thresholding at 130 HU. For each candidate voxel an image patch centered around the voxel is fed into a deep learning-based classification algorithm. The deep learning model has two components: a convolutional neural network, which takes the image patch and a precomputed coronary territory map as inputs, and a dense neural network which operates on the coordinates of the voxel. A final prediction is made by combining features from both components to determine whether the voxel belongs to the coronary arteries. The algorithm has been trained on over 1,200 ECG-gated calcium scoring scans and fine-tuned on over 550 chest CTs. Additional details on the computational pipeline and the network topology have been described by Chamberlin et al. [9].

The **total volume V** of the detected coronary calcium is used for threshold-based categorization. Several thresholds for total calcium volume have been proposed in the literature. For instance, based on the NELSON study, Mets et al. [16] showed that a coronary calcium volume of 100 mm<sup>3</sup> corresponds to an 8% increased risk of cardiovascular events and 500 mm<sup>3</sup> to an increased risk of 48%. These volumes were used as default thresholds in AI-Rad Companion Chest CT<sup>1</sup>. A third threshold at 10 mm<sup>3</sup> is used to compensate for image noise. Slightly different values have been used in a publication by van Assen et al [17]. An overview is provided in Table 3.

Category	Thresholds derived from Mets et al. [16]	Thresholds used by van Assen et al. [17]
I	$V < 10$	$V < 5$
II	$10 \leq V < 100$	$5 \leq V < 250$
III	$100 \leq V < 500$	$250 \leq V < 1000$
IV	$V \geq 500$	$V \geq 1000$

**Table 3:** Thresholds for coronary calcium volume V in mm<sup>3</sup>.

## Aorta diameter measurements

The aorta analysis pipeline consists of three steps: landmark detection, aorta segmentation, and diameter measurements.

Six aortic landmarks (Aortic Root, Aortic Arch Center, Brachiocephalic Artery Bifurcation, Left Common Carotid Artery, Left Subclavian Artery, and Celiac Trunk) are detected automatically based on Deep Reinforcement Learning [10].

The aortic root is used to define a ROI for the segmentation algorithm. Within the ROI the segmentation is performed using an adversarial DI2IN in a symmetric convolutional encoder-decoder architecture [11]. The front part is a convolutional encoder-decoder network with feature concatenation, and the backend is deep supervision network through multi-level. Blocks inside DI2IN consist of convolutional and upscaling layers. The algorithm has been trained on over 1,000 CT data sets including native and contrast-enhanced scans.

<sup>1</sup> In the software version available in the United States no default values are provided.

<sup>2</sup> In the software version available in the United States the thresholds cannot be adapted by the user.

Given the aorta mask, a centerline model is used to generate the aortic centerline. The centerline is used in combination with aortic landmarks to identify measurement planes at nine locations according to the guidelines of the American Heart Association [2]. In each of the measurement planes, multiple diameters are computed by computing intersections of rays starting from the centerline with the aortic mask. Based on these diameters, the maximum in-plane diameter and the diameter along a direction orthogonal to the maximal diameter are reported.

The **maximum diameters d** are used for threshold-based categorization. Recommended thresholds<sup>2</sup> for the severity of a potential aortic dilation or aneurysm have been derived from the AHA-Guidelines and values for the population mean and **standard variation (std)** given therein:

Category	Condition
I	$d \leq \text{mean} + 2 * \text{std}$
II	$d > \text{mean} + 2 * \text{std}$
III	$d > 1.5 * \text{mean}$
IV	$d \geq 5.5 \text{ cm}$

**Table 4:** Thresholds for aortic diameters d.

The aorta module works for both native and contrast-enhanced data with and without ECG-gating.

## Vertebra labeling and density measurement

The twelve thoracic vertebrae are localized and labeled using an algorithm based on wavelet features, AdaBoost, and local geometry constraints [18]. Around each vertebra center cylindrical regions of interest are used to measure the average HU-density of the trabecular bone.

## Vertebra segmentation and height measurement

The vertebra centers are also used to determine ROIs for the vertebra segmentation. Within the ROI the segmentation is performed using a DI2IN in a symmetric convolutional encoder-decoder architecture [11]. The algorithm has been trained on over 4,300 thoracic vertebrae.

From the segmentation masks the sagittal midplane is extracted and within this plane height measurements at anterior, medial, and posterior location. Afterwards, the **height ratios hr** are computed by comparing heights of neighboring vertebrae using the Genant severity grading method [19]. Although originally developed on chest radiographs, the Genant method is a widely used also in CT imaging [6]:

Category	Condition
I	$hr \geq 80\%$
II	$80\% > hr \geq 75\%$
III	$75\% > hr \geq 60\%$
IV	$hr < 60\%$

**Table 5:** Thresholds for vertebra height ratios hr.

## Data requirements

### Technical requirements

AI-Rad Companion Chest CT uses a single DICOM series as input for all modules. In general, the algorithms are intended to work with any chest CT series. However, there are a couple of technical properties required for the device to process the cases:

- Primary axial images (image orientation 1\0\0\0\1\1\0)
- Volume scans without gaps, no gantry tilt
- Slice thickness  $\leq 3$  mm (for MSK  $\leq 2$  mm), preferably  $\leq 1$  mm, see recommendations below
- Matrix size 512 x 512
- Photometric interpretation: MONOCHROME 2
- 16 bit, no lossy compression, samples per pixel: 1
- Rescale slope  $\leq 5$

The cardiac module (heart segmentation and coronary calcium detection) has the additional requirements that the images are without contrast enhancement and  $kVp \geq 100$ . That is because the initial candidate

generation step is based on HU-thresholding and the threshold is not valid for contrast-enhanced scans nor for  $kVp < 100$ . The topic has been discussed in detail by Vonder et al. [20] and in a corresponding Siemens Healthineers Whitepaper on calcium quantification on dedicated cardiac CT data [21].

### Scan Parameter Recommendations

Besides the coronary calcium detection, HU-thresholding is also used in the LAV-analysis of the lung parenchyma. As a consequence, the results of these two features are sensitive to image noise. Image noise in CT data depends on many parameters, most prominently on slice thickness, reconstruction kernel, and dose. Hence the combination of thin slices, hard kernels, and low dose may result in very noisy images. For such data the cardiac module would reject the case (if there are too many calcium candidates) and the LAV analysis may be confounded by noise-related LAV-patches [22].

Reconstruction kernel	Soft to medium kernel			Hard kernel		
	$\leq 1$ mm	1–2 mm	2–3 mm	$\leq 1$ mm	1–2 mm	2–3 mm
Lung nodules	●	●	●	●	●	●
Lung Parenchyma (LAV and opacities)	●	●	●	●	●	●
Aorta	●	●	●	●	●	●
Heart and Coronaries	●	●	●	●	●	●
Vertebrae	●	●	●	●	●	●

**Table 6:** Recommended scan parameters for AI-Rad Companion Chest CT.

● fully supported ● supported but results might be suboptimal ● not supported

Publication	Patient cohort	Feature(s) studied	Study size	Scanner Model(s)	Scan parameters
Chamberlin et al. [9]	Lung cancer screening	Lung nodules, cor. calcium	117	SOMATOM go.Top, Definition AS+, Definition Flash, and Force	protocol according to ACR-STR LDCT guidelines. slice thickness: 1.0 mm
van Assen et al. [17]	Paired Cardiac and Chest CTs, consecutive Chest CTs	Cor. calcium	95 + 168	SOMATOM Definition Flash, Definition AS+, and Force	slice thickness: 1.0 mm–3.0 mm, medium sharp kernel
Fischer et al. [22]	Emphysema	LAV	141	SOMATOM Definition Flash, Force, and Emotion	slice thickness 1.5 mm, comparing two kernels: lung (B60s) and soft tissue (B31s)
Yacoub et al. [23]	Consecutive cases	all	100	SOMATOM Definition Flash, and Force	slice thickness: 1.0 mm, soft tissue kernel
Rückel et al. [24]	Emergency CT	Lung nodules, aorta diam., cor. calcium, heart size, vert. heights	105	SOMATOM Force	slice thickness: 0.75 mm, soft tissue kernel Br36d
Fischer et al. [25]	COPD	Lung lobes, LAV	137	SOMATOM Definition Flash, Force, and Emotion	slice thickness 1.5 mm, lung kernel
Rückel et al. [27]	Aortic aneurysm follow-up	Aorta diam.	18 x 2	SOMATOM Definition Flash, Force, and Definition AS+, GE Optima CT660, Discovery 750 HD	slice thickness: 0.6 mm –3.0 mm, soft tissue kernel
Weikert et al. [28]	COVID-19 patients	Lobe volume, PO, PHO, LSS, LAV, heart size, cor. calcium, aorta diam	120	SOMATOM Definition AS+, and Force	slice thickness: 1.0 mm, soft tissue kernel
Homayounieh et al. [29]	COVID-19 patients	Lobe volume, PO, PHO	241	SOMATOM Definition Flash, Force, and Definition Edge, Emotion 16, GE Discovery 750 HD	slice thickness: 1.0 mm–2.0 mm, soft tissue kernel B20f
Abadia et al. [30]	Lung nodules in cases w/ complex lung disease	Lung lobes, Lung nodules	103 + 40	SOMATOM Force	slice thickness: 1.0 mm, sharp body kernel

**Table 7:** Scan parameters used in various publications using AI-Rad Companion Chest CT.

On the other hand, thin slices, i.e. high spatial resolution in z-direction, are required for most of the algorithms, in particular for accurate vertebrae height measurements (ideally slice thickness should be  $\leq 1$  mm), detailed delineation of lung lobes and accurate lung nodule volumetry.

In summary, Table 6 displays the recommendations of scan parameters for the individual modules of AI-Rad

Companion Chest CT. To achieve optimal results for all modules, it is recommended to use a thin slice with a soft to medium kernel.

Table 7 summarizes scan parameters used in various clinical studies using AI-Rad Companion Chest CT. Details about these and other studies will be discussed in the subsequent section.

## Considerations regarding patient population

Table 7 also illustrates that AI-Rad Companion Chest CT has been used to analyze a broad spectrum of patient cohorts:

- low dose lung cancer screening [9]
- consecutive cohorts, independent of particular clinical indications [23; 17] or with an indication unrelated to the features of AI-Rad Companion Chest CT like data from the emergency department [24]
- patients with known disease patterns relevant for the feature of investigation, like emphysema/COPD [25; 22], osteoporosis [26], aortic aneurysms [27] COVID-19 [28; 29]
- patients with known comorbidities that make the assessment of the feature under investigation more challenging, such as the detection of lung nodules in the presence of, e.g. ILD [30].

The broad spectrum illustrates the versatile and generic design of the algorithms of AI-Rad Companion Chest CT. On the other hand, one would always find cases where – due to severe pathology, comorbidity, or anatomical

deviation, but also due to imaging artefacts like motion or noise – one or more algorithms might fail or produce incorrect result. In that context it is also important to note that the output images generated by AI-Rad Companion Chest CT are designed in a way that the user can easily verify the correctness of the result. An example would be the sagittal MPR of the spine, see Figure 4.



**Figure 4:** AI-Rad Companion Chest CT output: Sagittal view of the spine including height and density measurement. The incorrect height measurement for T1 (due to image noise) can easily be verified by the user.

## Proof points: performance and clinical value

AI-Rad Companion Chest CT offers advanced ways of workflow customization. By design, all results are presented in the form of an annotated axial series, a 3D rendering, and a concise summary table – enabling seamless integration into a variety of different reading workflows. Moreover, a DICOM Structured Report with measurement results is provided.

### Accuracy and clinical value

In the study by Chamberlin et al. [9] N=117 lung cancer screening exams were processed by AI-Rad Companion Chest CT and analyzed with respect to lung nodules and coronary calcium. The agreement with expert reader has been found excellent (Cohen’s kappa of lung nodule detection: 0.846, intraclass correlation coefficient of coronary calcium volume: 0.904) at a sensitivity of 100 %

and 92.9 % (presence of lung nodules and presence of coronary calcifications, respectively) and a specificity of 70.8 % and 96.0 %, respectively. The authors also use the results for predicting of lung cancer and major adverse cardiac events at 1-year follow-up yielding AUC-values of 0.942 and 0.911, respectively, emphasizing that “this information can be used to improve diagnostic ability, facilitate intervention, improve morbidity and mortality, and decrease healthcare costs”.

Focusing on the other end of the spectrum of patient cohorts, namely patients with complex lung disease such as ILD, COPD, bronchitis, edema, and pulmonary embolism, Abadia et al. [30] investigated the accuracy of the lung nodule detection and localization (N = 103 plus 40 controls). On a patient level AI-Rad Companion Chest CT showed a sensitivity of 89.4 % and a specificity of 82.5 %.

On the individual nodule level sensitivity was 67.7%, similar to the accuracy reported for experienced radiologists.

On an unspecific but representative patient population, i.e. N = 100 consecutive cases, Yacoub et al. [23] reported sensitivity and specificity of all features of AI-Rad Companion Chest CT, see Table 8.

	N positive cases	Sensitivity		Specificity	
		AI	Report	AI	Report
Lung nodules	83	92.8%	97.6%	<b>82.4%</b>	100%
Emphysema	31	80.6%	<b>74.2%</b>	<b>66.7%</b>	97.1%
Aortic dilation	27	96.3%	<b>25.9%</b>	<b>81.4%</b>	100%
Coronary Calcium	59	89.8%	<b>75.4%</b>	100%	94.9%
Vertebra compression	9	100%	100%	<b>63.7%</b>	100%

**Table 8:** Sensitivity and specificity of AI-Rad Companion Chest CT and radiological reports on N = 100 consecutive cases as reported by Yacoub et al. [23].

The authors state that “such results align well with the general recommendation to maximize sensitivity when AI is being used in radiology to detect abnormalities, even at the expense of lower specificity, in order to ensure that fewer abnormal findings are missed. Our findings illustrate that the use of AI for diagnostic reading provides rather a support tool which is not intended to replace the role of a radiologist.” They conclude that “incorporating AI support into radiology workflows can provide significant added value to clinical radiology reporting”.

The low sensitivity of the radiologists in particular among the incidental findings has also been studied by Rückel et al. [24] in the particular time-critical setting of emergency CT. The following abnormalities were missing in the original reports but identified by AI-Rad Companion Chest CT and confirmed by radiologists in a consecutive cohort of N = 105 whole-body emergency CTs:

- 23.8 % increased heart size
- 16.2 % coronary calcifications

- 32.4 % aortic ectasia
- 1.9 % actionable lung nodules
- 12.4 % vertebra fractures

The authors point out that “In particular, the integration of different specialized algorithms in a single software solution is promising to avoid clinically too narrow AI applications. But also, with regard to less urgent applications of medical imaging, it should also be mentioned that especially non-radiology clinicians might even take more benefit from AI-assisted image analysis compared to anyway well-trained radiologists, e.g., in clinical settings without 24/7 radiology coverage or long turn-around times for radiology reporting.”

Consecutive patients (N = 168) were also enrolled in a study on coronary calcium detection by van Assen et al. [17]. Here the coronary calcium volume computed by AI-Rad Companion Chest CT was compared against the calcium volume obtained from manual calcium scoring. The correlation was found excellent (logarithmic correlation coefficient 0.923). By applying volume-thresholds (see Table 3) the AI-computed calcium volume was categorized into no, mild, moderate, and severe. The categories were compared against qualitative visual rating by an experienced cardiac radiologist. Results are shown in Table 9. 82 % of all cases were correctly classified with all wrongly classified scans being attributed to an adjacent category.

Expert\AI	No	Mild	Moderate	Severe
No	<b>60</b>	6	0	0
Mild	7	<b>44</b>	0	0
Moderate	0	6	<b>14</b>	5
Severe	0	0	4	<b>20</b>

**Table 9:** Category agreement between manual qualitative assessment and AI determined calcium volume as reported by van Assen et al. [17].

In a second arm of the study, N = 95 patients were identified which underwent both dedicated coronary calcium scoring exams (non-contrasted, ECG-gated cardiac CTs) and chest CTs within 1.5 years. For those patients, conventional calcium scoring was performed according to Agatston on the cardiac CTs and compared to the calcium volume computed by AI-Rad Companion Chest CT on the chest CT data. By design, the agreement of these results will be lower, simply because the data

compared originates from different acquisitions from different time points. Nevertheless, the correlation between manual Agatston score and calcium volume computed by AI-Rad Companion Chest CT was found excellent (logarithmic correlation coefficient 0.921). When comparing threshold-based categories (volume threshold as in Table 3 vs. standard Agatston risk categories), 70% of all cases were classified correctly, in only 5% the prediction was more than one category off. Moreover, a misclassification into the “no calcium” category, which – according to the authors [17] – “would have the largest impact on patient treatment, since these patients will be considered to have no/little cardiac risk”, occurred only in 3% of the cases.

Particular features of AI-Rad Companion Chest CT were also studied by Savage et al. [26], correlating the average HU-density of the vertebrae computed by the software with T-scores obtained from dual-energy X-ray absorptiometry (DEXA) on N= 65 patients yielding significant difference between healthy and osteoporotic (i.e.  $T < -2.5$ ) patients. This is supported by work by Cohen et al. [8] using manual HU-measurements. The authors found that a threshold of 110 HU could be used to identify osteoporotic patients with a specificity of 93%.

Two publications by Fischer et al. [25; 22] study the results of the lung lobe-based LAV analysis in emphysema/COPD patients (N= 141 and N= 137, respectively). The correlation of LAV with spirometry-based Tiffeneau index was -0.86, and 0.88 with GOLD stages, respectively. The LAV of the upper lobes “was also able to most clearly distinguish mild and moderate forms of COPD. This is particularly relevant due to the fact that early disease processes often elude conventional pulmonary function diagnostics. Earlier detection of COPD is a crucial element for positively altering the course of disease progression through various therapeutic options” [25].

In the course of the Covid-19 pandemic two papers by Weikert et al. [28] and Homayounieh et al. [29] investigated the use of AI-Rad Companion Chest CT features for the prediction of patient management and patient outcome in COVID-19 patients: Homayounieh et al. [29] used a combination of lung lobe volumes, PO and PHO yielding a “higher AUC for predicting ICU admission than subjective severity scores” (N= 241). Weikert et al. [28] added also cardiovascular metrics obtained from AI-Rad Companion Chest CT, namely heart volume, coronary calcium volume and aortic diameters, together with lab-findings yielding excellent predictions (AUC= 0.91,

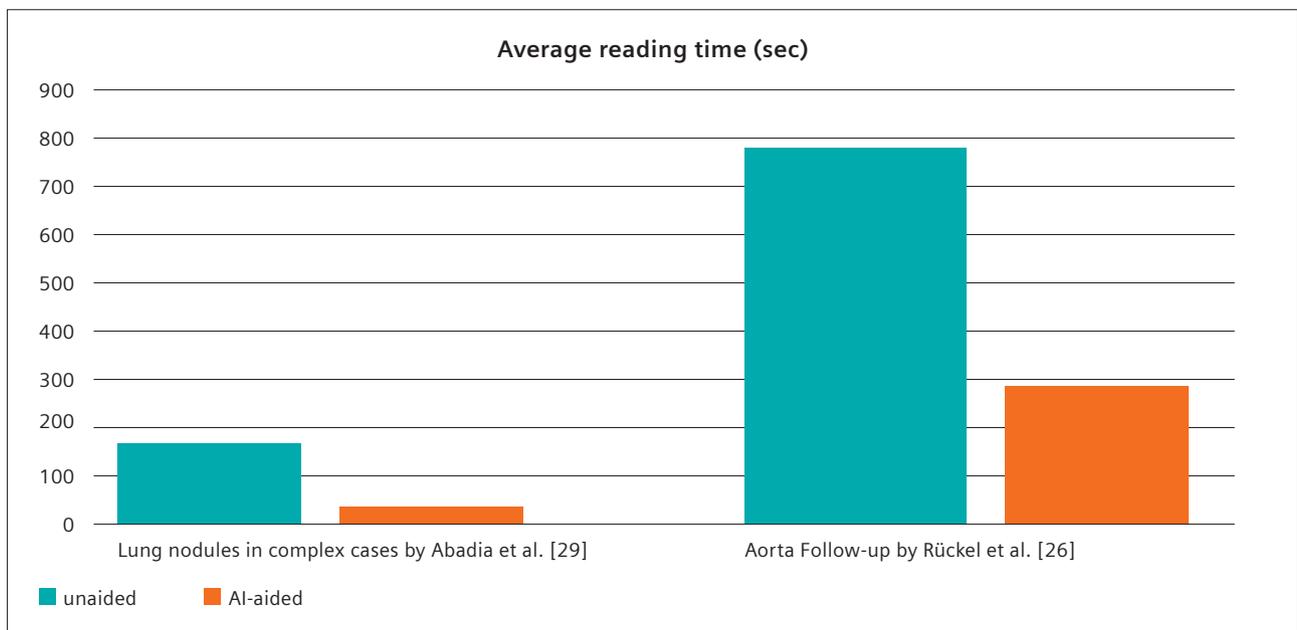


Figure 5: Average reading times with and without support of AI-Rad Companion Chest CT.

N = 120). In the work of Biebau et al [31], visual scores of lung injury were compared against AI-based scoring of the LSS on N = 182 consecutive Covid-19 patients yielding a very good correlation of 0.89.

## Efficiency and standardization

Increasing efficiency of the radiological workflow is key to manage increasing workload and at the same time saving healthcare cost. In the aforementioned study by Abadia et al. [30] on patients with complex lung disease average reading time for lung nodules was 2:44 min ± 0:54 min without support of AI-Rad Companion Chest CT. After a month of washout-period a random subset of N = 20 patients of the original study were reevaluated with support of AI-Rad Companion Chest CT. Here average reading time was reduced to 0:36 min, i.e. a significant reduction by 78%. Moreover, “the expert reported increased confidence for lung nodule detection for all 20 cases” [30].

Average reading time was also in the focus in a study by Rückel et al. [27] on N = 18 patients with aortic ectasia undergoing follow-up assessments (two timepoints per patient). Reading of the two time-points studies was performed by three radiologists with and without support of AI-Rad Companion Chest CT. Average reading time was reduced from 13:01 min to 4:46 min corresponding to a significant reduction by 63%. In addition, AI assistance reduced total diameter inter-reader variability by 42.5%. Figure 5 summarizes time savings reported by the various studies.

## Standalone performance

Besides the validation of AI-Rad Companion Chest CT in studies performed by academic sites, internal standalone performance tests on the individual algorithms have been performed:

- **Lung nodule detection:** For nodule size range of 4 to 30 mm sensitivity was 93.1% at 1 false positives per case (median), N = 316.
- **Lung lobe segmentation:** Average DICE coefficients for the individual lung lobes ranged between 0.95 and 0.98. Mean surface distance ranged between 0.5 mm and 1.0 mm. Volume error was between 1.5% and 3.5%. N = 4,500.

- **Opacity quantification:** Opaque regions were detected with a sensitivity of 89.4% at 0.544 average false positives per case. Correlation coefficient for PO was 0.945. 95%-Limits of agreement (LoA) of manual measurements of PO per lobe by two radiologists was established at 15.8%. Ratio of automatic PO measurements lying within the LoA was 93.0%. N = 149.
- **Heart segmentation:** Average DICE coefficient was 0.93. N = 274.
- **Coronary calcium detection:** Logarithmic correlation coefficient of total coronary calcium volume was 0.96. N = 381.
- **Aorta diameters:** Average absolute error in aorta diameters was 1.6 mm across all nine measurement locations and varied between 0.9 mm and 2.4 mm per location. N = 193.
- **Vertebra HU-density:** 95%-Limits of agreement (LoA) of manual density measurements by four radiologists was established at 64.1 HU. Ratio of automatic vertebra density measurements lying within the LoA was 98.8%. N = 150.
- **Vertebrae heights:** LoA of manual height measurements by four radiologists was established at 2.86 mm for slice thickness ≤ 1.0 mm, and at 3.20 mm for slice thickness > 1.0 mm, respectively. Ratio of automatic vertebra height measurements lying within the LoA was 95.1% for slice thickness ≤ 1.0 mm and 87.5% for slice thickness > 1.0 mm. N = 150.

## References

- 1 Guidelines for Management of Incidental Pulmonary Nodules Detected on CT Images: From the Fleischner Society 2017. MacMahon, H., Naidich, D. P., Goo, J. M., Lee, K. S., Leung, A., Mayo, J. R., Mehta, A. C., Ohno, Y., Powell, C. A., Prokop, M., Rubin, G. D., Schaefer-Prokop, C. M., Travis, W. D., Van Schil, P. E., & Bankier, A. A. 1, 2017, Radiology, Vol. 284, pp. 228-243.

- 2 2010 ACCF/AHA/AATS/ACR/ASA/SCA/SCAI/SIR/STS/SVM guidelines for the diagnosis and management of patients with Thoracic Aortic Disease. Hiratzka, L. F., et al. 13, 2010, *Circulation*, Vol. 121, pp. 266-369.
- 3 Managing Incidental Findings on Thoracic CT: Mediastinal and Cardiovascular Findings. A White Paper of the ACR Incidental Findings Committee. Munden, R. F., Carter, B. W., Chiles, C., MacMahon, H., Black, W. C., Ko, J. P., McAdams, H. P., Rossi, S. E., Leung, A. N., Boiselle, P. M., Kent, M. S., Brown, K., Dyer, D. S., Hartman, T. E., Goodman, E. M., Naidich, D. P., Kazerooni, E. A., Berland, L. 8, 2018, *J Am Coll Radiol.*, Vol. 16, pp. 1097-1096.
- 4 2016 SCCT/STR guidelines for coronary artery calcium scoring of noncontrast noncardiac chest CT scans: A report of the Society of Cardiovascular Computed Tomography and Society of Thoracic Radiology. Hecht, H. S., Cronin, P., Blaha, M. J., Budoff, M. J., Kazerooni, E. A., Narula, J., Yankelevitz, D., & Abbara, S. 5, 2017, *J Thorac Imaging*, Vol. 32, pp. W54-W66.
- 5 Reporting incidental coronary, aortic valve and cardiac calcification on non-gated thoracic computed tomography, a consensus statement from the BSCI/ BSCCT and BSTI. Williams, M. C., Abbas, A., Tirr, E., Alam, S., Nicol, E., Shambrook, J., Schmitt, M., Hughes, G. M., Stirrup, J., Holloway, B., Gopalan, D., Deshpande, A., Weir-McCall, J., Agrawal, B., Rodrigues, J., Brady, A., Roditi, G., Robinson, G., & Bull, R. 1117, 2021, *Br J Radiol*, Vol. 94, p. 20200894.
- 6 Adams, J.E., Lenchik, L., Roux, C., & Genant, H. K. Radiological Assessment of Vertebral Fracture. International Osteoporosis Foundation Vertebral Fracture Initiative Resource Document Part II. 2010, pp. 1-49.
- 7 Opportunistic screening for osteoporosis using abdominal computed tomography scans obtained for other indications. *Annals of internal medicine*. Pickhardt, P. J., Pooler, B. D., Lauder, T., del Rio, A. M., Bruce, R. J., & Binkley, N. 8, 2013, *Ann Intern Med*, Vol. 158, pp. 588-595.
- 8 Opportunistic screening for osteoporosis and osteopenia by routine computed tomography scan: A heterogeneous, multiethnic, middle-eastern population validation study. Cohen, A., Foldes, A. J., Hiller, N., Simanovsky, N., & Szalat, A. 2021, *Eur J Radiol*, Vol. 136, p. 109568.
- 9 Automated detection of lung nodules and coronary artery calcium using artificial intelligence on low-dose CT scans for lung cancer screening: accuracy and prognostic value. Chamberlin, J., Kocher, M. R., Waltz, J., Snoddy, M., Stringer, N., Stephenson, J., Sahbaee, P., Sharma, P., Rapaka, S., Schoepf, U. J., Abadia, A. F., Sperl, J., Hoelzer, P., Mercer, M., Somayaji, N., Aquino, G., & Burt, J. R. 2021, *BMC Medicine*, Vol. 19, p. 55.
- 10 MultiScale Deep Reinforcement Learning for Real-Time 3D-Landmark Detection in CT Scans. Ghesu, F. C., Georgescu, B., Zheng, Y., Grbic, S., Maier, A., Hornegger, J., & Comaniciu, D. 1, 2019, *IEEE Trans Pattern Anal Mach Intell.*, Vol. 41, pp. 176-189.
- 11 Automatic Liver Segmentation Using an Adversarial Image-to-Image Network. Yang, D., Xu, D., Zhou, S. K., Georgescu, B., Chen, M., Grbic, S., Metaxas, D., & Comaniciu, D. 2017. *MICCAI*. pp. 507-515.
- 12 Quantifying the extent of emphysema: factors associated with radiologists' estimations and quantitative indices of emphysema severity using the ECLIPSE cohort. Gietema, H. A., Müller, N. L., Fauerbach, P. V., Sharma, S., Edwards, L. D., Camp, P. G., & Coxson, H. O. 6, 2011, *Academic radiology*, Vol. 18, pp. 661-671.
- 13 U-Net: Convolutional Networks for Biomedical Image Segmentation. Ronneberger, O., Fischer, P., & Brox, T. 2015. *MICCAI*. pp. 234-241.
- 14 Automated Quantification of CT Patterns Associated with COVID-19 from Chest CT. Chaganti, S., Balachandran, A., Chabin, G., Cohen, S., Flohr, T., Georgescu, B., Grenier, P., Grbic, S., Liu, S., Mellot, F., Murray, N., Nicolaou, S., Parker, W., Re, T., Sanelli, P., Sauter, A. W., Xu, Z., Yoo, Y., Ziebandt, V., & Comaniciu, D. 4, 2020, *Radiology: Artificial Intelligence*, Vol. 2.
- 15 Chest CT Findings in Coronavirus Disease-19 (COVID-19): Relationship to Duration of Infection. Bernheim, A., Mei, X., Huang, M., Yang, Y., Fayad, Z. A., Zhang, N., Diao, K., Lin, B., Zhu, X., Li, K., Li, S., Shan, H., Jacobi, A., & Chung, M. 3, 2020, *Radiology*, Vol. 295, p. 200463.
- 16 Lung cancer screening CT-based prediction of cardiovascular events. Mets, O. M., Vliegenthart, R., Gondrie, M. J., Viergever, M. A., Oudkerk, M., de Koning, H. J., Mali, W. P., Prokop, M., van Klaveren, R. J., van der Graaf, Y., Buckens, C. F., Zanen, P., Lammers, J. W., Groen, H. J., Isgum, I., & de Jong, P. A. 8, 2013, *JACC: Cardiovascular Imaging*, Vol. 6, pp. 899-907.

- 17 Automatic coronary calcium scoring in chest CT using a deep neural network in direct comparison with non-contrast cardiac CT: A validation study. van Assen, M., Martin, S. S., Varga-Szemes, A., Rapaka, S., Cimen, S., Sharma, P., Sahbaee, P., De Cecco, C. N., Vliegenthart, R., Leonard, T. J., Burt, J. R., & Schoepf, U. J. 2021, *Eur J Radiol*, Vol. 134, p. 109428.
- 18 Zhan, Y., Jian, B., Maneesh, D., & Zhou, X.S. Cross-Modality Vertebrae Localization and Labeling Using Learning- Based Approaches. [ed.] S. Li and J. Yao. *Spinal Imaging and Image Analysis. Lecture Notes in Computational Vision and Biomechanics*. s.l.: Springer, 2015, Vol. 18.
- 19 Vertebral fracture assessment using a semiquantitative technique. Genant, H. K., Wu, C. Y., van Kuijk, C., & Nevitt, M. C. 9, 1993, *J Bone Miner Res.*, Vol. 8, pp. 1137-48.
- 20 The impact of dose reduction on the quantification of coronary artery calcifications and risk categorization: A systematic review. Vonder, M., van der Werf, N. R., Leiner, T., Greuter, M., Fleischmann, D., Vliegenthart, R., Oudkerk, M., & Willeminck, M. J. 5, 2018, *J Cardiovasc Comput Tomogr*, Vol. 12, pp. 352-363.
- 21 Siemens Healthineers. Agatston Score calcium quantification with arbitrary tube voltage. 2018.
- 22 Comparison of Artificial Intelligence-Based Fully Automatic Chest CT Emphysema Quantification to Pulmonary Function Testing. Fischer, A. M., Varga-Szemes, A., van Assen, M., Griffith, L. P., Sahbaee, P., Sperl, J. I., Nance, J. W., & Schoepf, U. J. 5, 2020, *Am J Roentgenol*, Vol. 214, pp. 1065-1071.
- 23 Performance of an Artificial Intelligence-Based Platform Against Clinical Radiology Reports for the Evaluation of Noncontrast Chest CT. Yacoub, B., Kabakus, I. M., Schoepf, U. J., Giovagnoli, V. M., Fischer, A. M., Wichmann, J. L., Martinez, J. D., Sharma, P., Rapaka, S., Sahbaee, P., Hoelzer, P., Burt, J. R., Varga-Szemes, A., & Emrich, T. 21, 2021, *Acad. Radiol.*, Vols. S1076-6332, pp. 00070-2.
- 24 Reduction of missed thoracic findings in emergency whole-body. Rueckel, J., Sperl, J., Kaestle, S., Hoppe, B., Fink, N., Rudolph, J., Schwarze, V., Geyer, T., Strobl, F., Ricke, J., Ingrisich, M., & Sabel, B. 6, 2021, *Quantitative Imaging in Medicine and Surgery*, Vol. 11, pp. 2486-2498.
- 25 Artificial Intelligence-based Fully Automated Per Lobe Segmentation and Emphysema-quantification Based on Chest Computed Tomography Compared With Global Initiative for Chronic Obstructive Lung Disease Severity of Smokers. Fischer, A. M., Varga-Szemes, A., Martin, S. S., Sperl, J. I., Sahbaee, P., Neumann, D., Gawlitza, J., Henzler, T., Johnson, C. M., Nance, J. W., Schoenberg, S. O., & Schoepf, U. J. *Suppl 1*, 2020, *J Thorac Imaging*, Vol. 35, pp. S23-S34.
- 26 Utilizing Artificial Intelligence to Determine Bone Mineral Density Via Chest Computed Tomography. Savage, R. H., van Assen, M., Martin, S. S., Sahbaee, P., Griffith, L. P., Giovagnoli, D., Sperl, J. I., Hopfgartner, C., Kärgel, R., & Schoepf, U. J. 2020 : s.n., *J Thorac Imaging*, Vol. 35, pp. S35-S39.
- 27 Artificial intelligence assistance improves reporting efficiency of thoracic aortic aneurysm CT follow-up. Rueckel, J., Reidler, P., Fink, N., Sperl, J., Geyer, T., Fabritius, M. P., Ricke, J., Ingrisich, M., & Sabel, B. O. 2021, *Eur J Radiol*, Vol. 134, p. 109424.
- 28 Prediction of Patient Management in COVID-19 Using Deep Learning-Based Fully Automated Extraction of Cardiothoracic CT Metrics and Laboratory Findings. Weikert, T., Rapaka, S., Grbic, S., Re, T., Chaganti, S., Winkel, D. J., Anastasopoulos, C., Niemann, T., Wiggli, B. J., Bremerich, J., Twerenbold, R., Sommer, G., Comaniciu, D., & Sauter, A. W. 2021, *Korean J Radiol*, Vol. 10.3348/kjr.2020.0994.
- 29 Multicenter Assessment of CT Pneumonia Analysis Prototype for Predicting Disease Severity and Patient Outcome. Homayounieh, F., Bezerra Cavalcanti Rockenbach, M. A., Ebrahimi, S., Doda Khera, R., Bizzo, B. C., Buch, V., Babaei, R., Karimi Mobin, H., Mohseni, I., Mitschke, M., Zimmermann, M., Durlak, F., Rauch, F., Digumarthy, S. R., & Kalra, M. K. 2021, *J Digit Imaging*, Vols. doi: 10.1007/s10278-021-00430-9.
- 30 Diagnostic accuracy and performance of artificial intelligence in detecting lung nodules in patients with complex lung disease: a non-inferiority study. Abadia, A.F. ACCEPTED IN JTI - IN PRESS - TO BE UPDATED.
- 31 Comparing Visual Scoring of Lung Injury with a Quantifying AI-Based Scoring in Patients with COVID-19. Biebau, C., Dubbeldam, A., Cockmartin, L., Coudyzer, W., Coolen, J., Verschakelen, J. & De Wever, W. 1, 2021, *Journal of the Belgian Society of Radiology*, Vol. 105, p. 16.

AI-Rad Companion Chest CT is not commercially available in all countries, and its future availability cannot be ensured.

The information in this document contains general technical descriptions of specifications and options as well as standard and optional features which do not always have to be present in individual cases, and which may not be commercially available in all countries. Due to regulatory reasons their future availability cannot be guaranteed. Please contact your local Siemens organization for further details.

Siemens reserves the right to modify the design, packaging, specifications, and options described herein without prior notice. Please contact your local Siemens sales representative for the most current information.

Note: Any technical data contained in this document may vary within defined tolerances. Original images always lose a certain amount of detail when reproduced.

---

**Siemens Healthineers Headquarters**

Siemens Healthcare GmbH  
Henkestr. 127  
91052 Erlangen, Germany  
Phone: +49 9131 84-0  
siemens-healthineers.com