Present Limitations of Artificial Intelligence in the Emergency Setting – Performance Study of a Commercial, Computer-Aided Detection Algorithm for Pulmonary Embolism

Grenzen künstlicher Intelligenz in der Notfallbefundung – eine Leistungsanalyse eines kommerziellen, computerbasierten Algorithmus zur Detektion von Lungenarterienembolien

Authors
Katharina Müller-Peltzer[1] 🆔, Lena Kretzschmar[2], Giovanna Negrão de Figueiredo[2], Alexander Crispin[3], Robert Stahl[4], Fabian Bamberg[1], Christoph Gregor Trumm[4]

Affiliations
1 Klinik für Diagnostische und Interventionelle Radiologie, Albert-Ludwigs-Universität Freiburg Medizinische Fakultät, Freiburg, Deutschland
2 Klinik und Poliklinik für Radiologie, Ludwig-Maximilians-Universität, München, Deutschland
3 Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie, Klinikum der Universität München-Großhadern, München, Deutschland
4 Institut für Diagnostische und Interventionelle Neuroradiologie, Klinikum der Universität München-Großhadern, München, Deutschland

Correspondence
Dr. Katharina Müller-Peltzer
Klinik für Diagnostische und Interventionelle Radiologie,
Albert-Ludwigs-Universität Freiburg, Hugstetter Str. 55,
79106 Freiburg im Breisgau, Germany
Tel.: +49/7 61/27 03 81 90
Katharina.Mueller-Peltzer@uniklinik-freiburg.de

ABSTRACT

Purpose Since artificial intelligence is transitioning from an experimental stage to clinical implementation, the aim of our study was to evaluate the performance of a commercial, computer-aided detection algorithm of computed tomography pulmonary angiograms regarding the presence of pulmonary embolism in the emergency room.

Materials and Methods This retrospective study includes all pulmonary computed tomography angiogram studies performed in a large emergency department over a period of 36 months that were analyzed by two radiologists experienced in emergency radiology to set a reference standard. Original reports and computer-aided detection results were compared regarding the detection of lobar, segmental, and subsegmental pulmonary embolism. All computer-aided detection findings were analyzed concerning the underlying pathology. False-positive findings were correlated to the contrast-to-noise ratio.

Results Expert reading revealed pulmonary embolism in 182 of 1229 patients (49 % men, 10–97 years) with a total of 504 emboli. The computer-aided detection algorithm reported 3331 findings, including 258 (8 %) true-positive findings and 3073 (92 %) false-positive findings. Computer-aided detection analysis showed a sensitivity of 47 % (95 %CI: 33–61 %) on the lobar level and 50 % (95 %CI 43–56 %) on the subsegmental level. On average, there were 2.25 false-positive findings per study (median 2, range 0–25). There was no significant correlation between the number of false-positive findings and the contrast-to-noise ratio (Spearman's Rank Correlation Coefficient = 0.09). Soft tissue (61.0 %) and pulmonary veins (24.1 %) were the most common underlying reasons for false-positive findings.

Conclusion Applied to a population at a large emergency room, the tested commercial computer-aided detection algorithm faced relevant performance challenges that need to be addressed in future development projects.

Key Points:

- Computed tomography pulmonary angiograms are frequently acquired in emergency radiology.
- Computer-aided detection algorithms (CADs) can support image analysis.
- CADs face challenges regarding false-positive and false-negative findings.
- Radiologists using CADs need to be aware of these limitations.
- Further software improvements are necessary ahead of implementation in the daily routine.

Citation Format

ZUSAMMENFASSUNG

Ziel Da zunehmend Anwendungen künstlicher Intelligenz im klinischen Alltag implementiert werden, war das Ziel unserer Studie, die Leistung eines kommerziellen, computerassistierten Detektionsalgorithmus für Lungenarterienembolien in Computertomografie-Pulmonalisangiografien im Rahmen einer Notfallbefundung zu analysieren.

Material und Methoden In diese retrospektive Studie wurden alle Computertomografie-Pulmonalisangiografien eingeschlossen, die über einen Zeitraum von 36 Monaten in einer großen deutschen Notaufnahme gefahren wurden. Alle eingeschlossenen Untersuchungen wurden von 2 in Notfallra-diologie erfahrenen Radiologen bewertet, um einen Referenzstandard festzulegen. Die Originalbefunde wurden mit den Ergebnissen des computerassistierten Detektionsalgorithmus in Bezug auf die Erkennung lobärer, segmentaler und subsegmentaler Lungenarterienembolien verglichen. Für alle Ergebnisse des computerbasierten Detektionsalgorithmus wurde die zugrunde liegende Pathologie definiert. Falsch positive Ergebnisse des Detektionsalgorithmus wurden mit dem Kontrast-Rausch-Verhältnis verglichen.

Ergebnisse Im Rahmen der Referenzbefundung wurden 504 Emboli in 182 von insgesamt 1229 Patienten (49 % männlich, 10–97 Jahre) detektiert. Der computerassistierte Detektionsalgorithmus zeigte insgesamt 3331 Ergebnisse an, hiervon waren 258 (8 %) richtig positiv und 3073 (92 %) falsch positiv. Die Sensitivität des Detektionsalgorithmus betrug 47 % (95 %-Konfidenzintervall 33–61 %) für lobäre und 50 % (95 %-Konfidenzintervall 43–56 %) für subsegmentale Lungenarterienembolien. Im Durchschnitt wurden 2,25 Ergebnisse pro Untersuchung (Median 2, 0–25) angezeigt. Es gab keine signifikante Korrelation zwischen der Anzahl der falsch positiven Ergebnisse und dem Kontrast-Rausch-Verhältnis (Spearman's Rangkorrelationskoeffizient = 0,09). Weichteilgewebe (61,0 %) und Pulmonalvenen (24,1 %) waren die häufigsten Ursachen für falsch positive Ergebnisse.

Schlussfolgerung Die Anwendung des getesteten computerassistierten Detektionsalgorithmus auf ein Patientenkollektiv einer großen Notaufnahme deckt die Schwachstellen des Algorithmus auf, die im Rahmen zukünftiger Projektentwicklungen optimiert werden sollten.

Introduction

Acute pulmonary embolism (PE) is a common disease with potentially high morbidity and mortality [1]. In international epidemiological studies, an annual incidence of acute PE of 39–115/100 000 people was calculated [2]. In recent years, an increasing incidence rate of acute pulmonary embolism has also been recorded. In particular, this increase can be attributed to the greater use of computed tomography (CT) for diagnostic workup in the emergency department resulting in the detection of even small peripheral pulmonary embolisms on the subsegmental level. In spite of scoring systems for estimating the probability of a PE, patients whose symptoms are due to other causes are also being increasingly examined [3–5]. Every computed tomography pulmonary angiography (CTPA) examination provides a significant amount of image data. In particular, given the growing number of CT examinations, exact evaluation of distal subsegments during image interpretation in an emergency setting is challenging. The goal of computer-aided detection algorithms (CADs) is to provide support for such time-consuming image analyses. While CAD systems are already well-established for the detection of pulmonary masses, PE-CADs are not yet used in the clinical routine. This can be the result of concern regarding false-positive CAD results or regarding the significant time required to use the program [6].

The goal of this study was to evaluate the use of a commercial computer-based algorithm for detecting pulmonary embolisms based on a patient population at a large German emergency department.

Materials and Methods

This monocentric retrospective study was approved by the local ethics committee. All CT examinations included in the study were clinically indicated.

Study population

All patients who underwent a CT examination in the period from 1/17/2017 to 1/17/2017 on an emergency CT unit at a large German university hospital due to suspicion of pulmonary embolism were included in the study. The corresponding datasets were identified by a keyword search of the examination names in the Picture Archiving and Communication System (PACS, Siemens Healthineers, Erlangen, Germany). The search yielded 2076 datasets for the indicated period. The flowchart in ► **Fig. 1** shows how
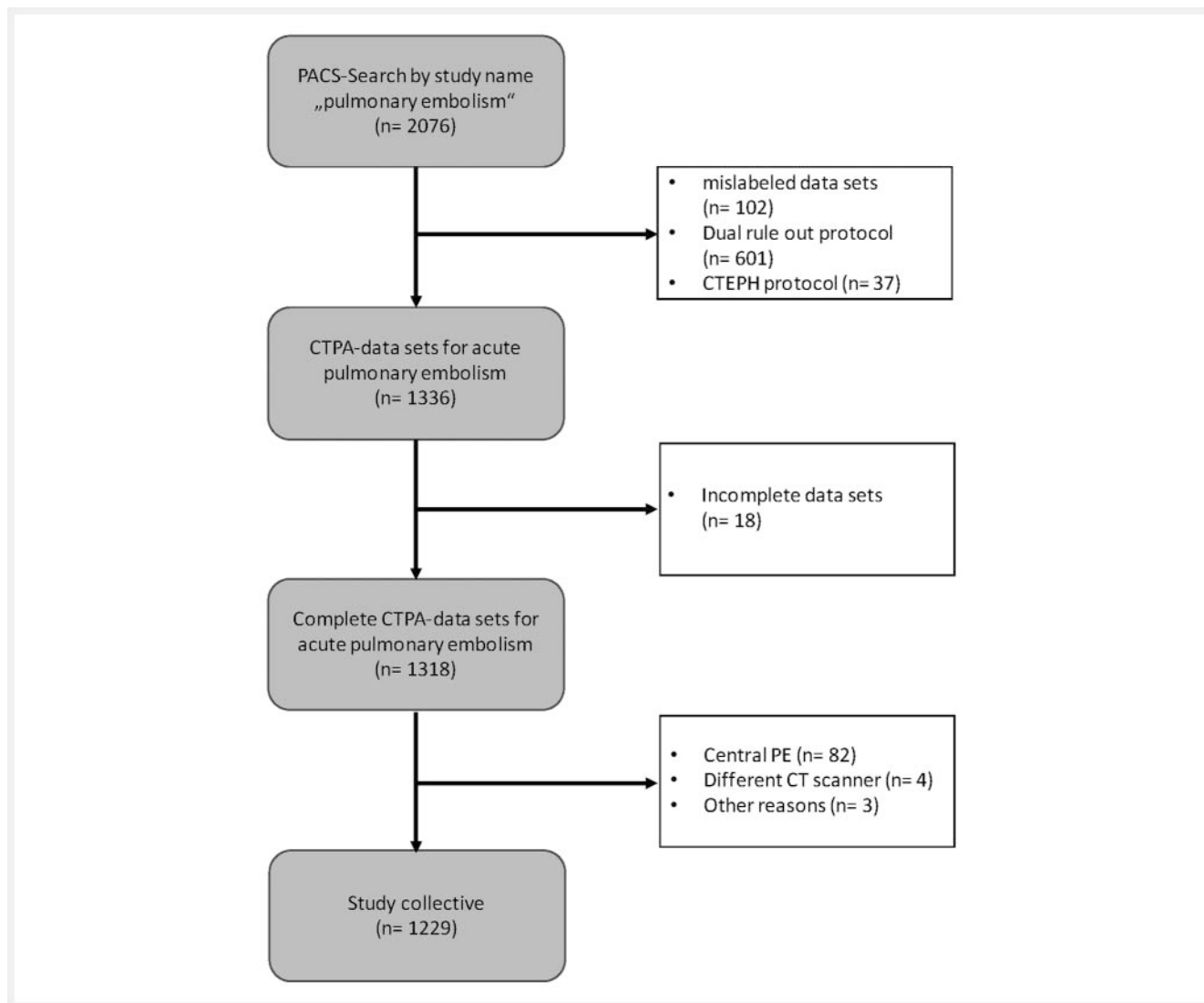
▶ **Fig. 1** Flowchart of the included data sets. PACS – picture archiving and communication system; CTEPH – chronic thromboembolic pulmonary hypertension; CTPA – computed tomography pulmonary angiograms; PE – pulmonary embolism; CT – computed tomography.

many datasets were excluded and the reason for exclusion. Incomplete datasets and those incorrectly defined as CTPA were excluded. Examinations acquired on another CT unit, acquired using the dual-energy technique to rule out chronic thromboembolic pulmonary hypertension (CTEPH), or acquired as a combination protocol ("dual rule out protocol") were also excluded. The latter could not be detected by the CAD software and thus could not be analyzed. One dataset was excluded in each case because of detection of a bone cement embolism, because of prior pneumonectomy, and because of prior Fontan procedure ("other reasons"). Since central pulmonary embolisms in the pulmonary arteries are a visual diagnosis, these datasets were also excluded from the study. The analyzed study population included 1229 patients.

CT acquisition

All CTPA datasets included in the study population were acquired on the Dual-Source *SOMATOM Force* CT unit (Siemens Healthineers, Erlangen, Germany) with a gantry rotation time of 0.25 s/ rotation, a pitch of 1.9, a collimation of 192×0.6 mm, and a slice thickness of 0.75 mm. The tube voltage and tube current were automatically adjusted using the settings "CAREkV" and "CARE-Dose". The reference points were a tube voltage of 100 kV and a tube current time product of 220 mAs. To optimize the contrast, bolus triggering was used with a region of interest (ROI) in the pulmonary trunk and an automated scan start 6 seconds after the density within the ROI exceeded 100 Hounsfield Units (HU). Contrast agent was injected with a flow rate of 5 ml/s. The standard amount of contrast agent that was administered was 50 ml followed by a 100-ml bolus of a 0.9 % saline solution. There was a deviation from the indicated amount of contrast agent in 117 cases with a documented BMI ≥ 30 kg/m². In these cases

70 ml of contrast agent were administered [7]. In the years 2014 and 2015, Imeron400 (Bracco Imaging GmbH, Constance, Germany) with an iodine concentration of 400 mg iodine/ml was used and in the years 2016 and 2017 Ultravist370 (Bayer Vital GmbH, Leverkusen, Germany) with an iodine concentration of 370 mg iodine/ml was used.

Image analysis

CAD analysis

Automated image analysis was performed with the software *syngo.via* (Siemens Healthineers, Erlangen, Germany) version VB20A as part of the study analysis. For PE-CAD analysis, the datasets were opened with the "CT Chest Pain + PE CAD". In the "CT Lung Vessels" module, automated analysis of the pulmonary arterial circulation was performed. PE-CAD systems use grayscale analysis to detect areas with reduced contrast enhancement adjacent to vessel segments with significant contrast enhancement [8]. The PE-CAD results referred to in the following as findings must then be assessed by the radiologist as false positive or false negative. False-positive findings can be deleted, while true-positive findings are confirmed and can be sent to the PACS following evaluation. The software does not evaluate the individual findings or the complete report. That is the responsibility of the radiologist.

The CAD analysis performed as part of this study was evaluated by the gold standard reader. All CAD findings were evaluated as true positive or false positive compared to the visual analysis performed during gold standard reading. The underlying cause of every false-positive result was defined in a consensus reading: Soft tissue, pulmonary vein, pneumonia, atelectasis, contrast inflow, drainage material, intrapulmonary calcification.

Gold standard

Visual image analysis was performed as part of the gold standard reading using the software *syngo.via* (Siemens Healthineers, Erlangen, Germany), version VB20A. All datasets from the study population (n = 1229) were visually evaluated by two radiologists with 5 and 15 years of emergency radiology experience as part of a consensus reading. This visual consensus reading was defined as the gold standard. The results of the CAD analysis were evaluated by the same radiologist via consensus reading and were classified as true positive or false positive compared to the gold standard. An embolism identified correctly multiple times by the CAD was classified as a true-positive result. Every embolism that was not identified by the CAD in comparison to the gold standard was classified as false negative. If a PE was not identified by the gold standard or the CAD, the examination was classified as true negative.

When reading a CTPA in an emergency setting, not each individual embolism but rather the most proximal embolism and its peripheral extension are described since providing a greater degree of detail does not affect the treatment decision. Therefore, the findings acquired in the emergency setting were compared to the gold standard on the patient level and classified as true positive, false positive, true negative, or false negative.

$$CNR = \frac{(HU_{\text{segmental artery}} - HU_{\text{muscles}})}{[(SD_{\text{segmental artery}} - SD_{\text{muscles}})/2]}$$

▶ **Fig. 2** Formula for calculating the contrast-to-noise ratio. CNR – contrast-to-noise ratio; HU – Hounsfield units; SD – standard deviation.

Routine reading at the emergency department

The radiological findings documented in the patient file and created during routine reading at the emergency department were analyzed visually with Siemens PACS Software (Siemens Healthineers, Erlangen, Germany) without the support of CAD software.

Definition of pulmonary embolism

An acute pulmonary embolism was defined as an intravascular central filling defect surrounded by a thin rim of contrast or complete occlusion of a focally dilated vessel [9]. Depending on the location, the PE was defined as lobar, segmental, or subsegmental and was assigned to a side. An embolism stretching continuously over multiple anatomical planes was assigned to the most central plane and was classified as an embolism [6].

Contrast-to-noise ratio

For the quantitative assessment of image quality, the contrast-to-noise ratio (CNR) in every dataset was determined. The measurements were performed in syngo.via. The equally sized ROIs were placed in the lumen of the left lower lobe superior segmental pulmonary artery and in the paravertebral musculature, and the HU values were applied with a standard deviation to the formula shown in ▶ **Fig. 2**. In the case of a filling defect in the left lower lobe apical segmental pulmonary artery, the ROI was placed in an adjacent lower lobe segmental pulmonary artery without a filling defect.

Statistical analysis

Continuous data are provided as mean with standard deviation and as a percentage. Categorical data are provided with absolute and relative frequency (as a percentage). The sensitivity and specificity are provided as a percentage with the corresponding 95 % confidence interval (CI). The dependence between the occurrence of segmental and subsegmental pulmonary artery embolisms was tested with the Chi-square test. The correlation between the number of CAD results and the CNR was examined using the Spearman's correlation analysis. A p-value < 0.05 was defined as significant. Excel 2016, Version 16.15 (Microsoft Corporation, Redmond, WA, USA) and Prism 8, Version 8.0.0 (GraphPad Software Inc., La Jolla, CA, USA) were used for the analysis.

Results

Patient characteristics

The datasets for 1229 patients were included in the study, including 607 (49.4 %) male patients and 622 (50.6 %) female patients 10 to 97 years old (mean 64 ± 17 years). The study population did not diverge significantly from the group in the initial database query (n = 2076) with 1062 (51.2 %) male patients and 1014 (48.8 %) female patients between 0 and 97 years old (mean 68 ± 18 years).

Gold standard results

A pulmonary embolism was detected in 182 of 1229 (14.8 %) patients. 93 (7.57 %) patients were male and 89 (7.24 %) were female. Multiple embolisms were detected in 106 (58.24 %) patients and a single pulmonary embolism was detected in 76 (41.76 %) patients. It total, 504 embolisms were described with 51 (10.12 %) being on the lobe level, 219 (43.45 %) on the segmental level, and 234 (46.43 %) on the subsegmental level. The majority of embolisms were right segmental (137 embolisms; 27.18 %) and right subsegmental (146 embolisms; 28.97). The Chi-square test confirmed a statistical dependence between segmental and subsegmental pulmonary artery embolisms (X^2 (1, N = 453) = 28.2, p < 0.0001).

CAD results

CAD analysis resulted in 3331 findings in 1229 datasets. These ranged from 0 to a maximum of 25 findings per dataset with a mean of 2.7 ± 2.61 findings. Among 182 patients with a PE detected by the gold standard, CAD analysis yielded at least one true-positive finding in 118 patients, resulting in a sensitivity of 64.8 % (95 % CI: 0.58–0.72). The embolisms identified by gold standard reading were missed by the CAD software in 64 patients. No finding was detected in 13 of these 64 datasets. Findings were identified in the remaining 51 datasets. However, they were false-positive findings. Among 1047 patients with no PE according to the gold standard, CAD analysis yielded at least one finding in 821 patients and no finding in 226 patients corresponding to a specificity of 21.6 % (95 % CI: 0.19–0.24), a positive predictive value (PPV) of 12.6 % (95 % CI: 0.11–0.15), and a negative predictive value (NPV) of 77.9 % (95 % CI: 0.73–0.82). The frequency distribution of the number of false-positive CAD findings per patient is shown in ▶ Fig. 3.

Among 504 embolisms identified by the gold standard, CAD-detected 258 embolisms (see ▶ Fig. 4) corresponding to a sensitivity of 51.2 % (95 % CI: 0.47–0.56). 24 of 51 embolisms (sensitivity: 47.1 %; 95 % CI: 0.33–0.61) were detected on the lobe level, 118 of 219 embolisms (sensitivity: 53.9 %; 95 % CI: 0.47–0.61) on the segmental level, and 116 of 234 embolisms (sensitivity: 49.6 %; 95 % CI: 0.43–0.56) on the subsegmental level.

Among the 3331 CAD findings, 3073 (92.3 %) were classified as false positive by the consensus reading (see ▶ Fig. 4). Therefore, on average, there were 2.5 ± 2.54 false-positive findings per patient. An underlying cause was assigned to every false-positive
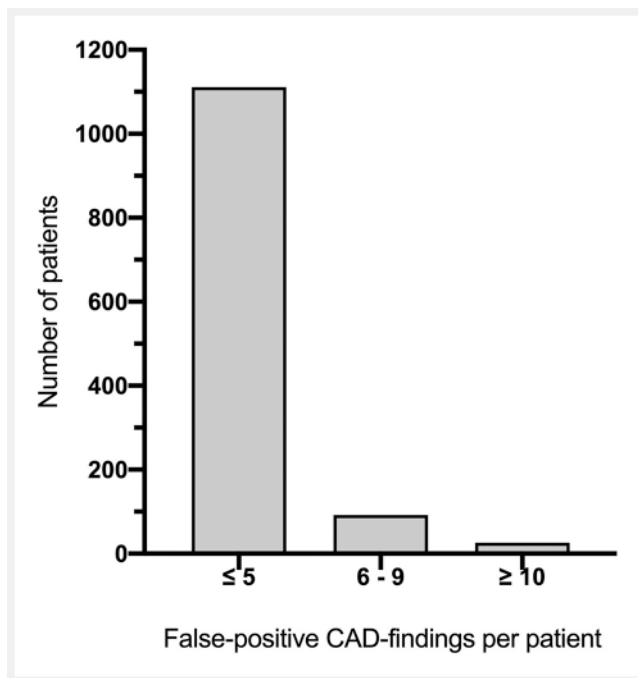
▶ Fig. 3 Graph of the frequency distribution of false-positive CAD findings. CAD - computer-aided detection.
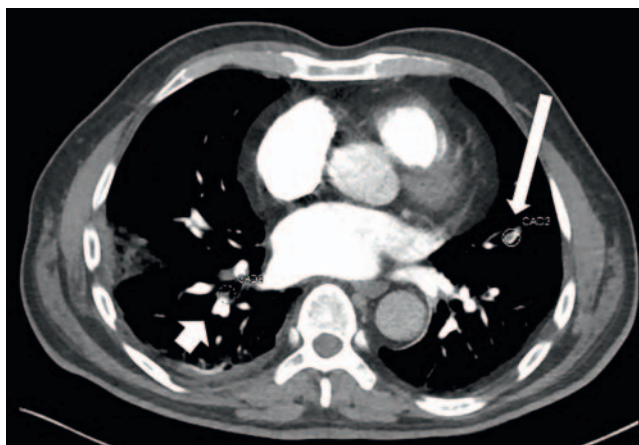
▶ Fig. 4 Examples of a true-positive (long, white arrow) and a false-positive (short, white arrow) finding.

CAD finding during the consensus reading. These are summarized in ▶ Table 1.

The Chi-square test confirmed a statistical correlation between the number of CAD findings and the PE status, (X^2 (1, N = 1229) = 20.6, p < 0.00 001). The contingency table including the column percentage is shown in ▶ Table 2.

Results of routine reading at the emergency department

Routine reading by the radiologist on duty described a pulmonary embolism in 180 of the 182 PE-positive datasets defined by the gold standard. This corresponds to a sensitivity for routine reading

at the emergency department compared to gold standard reading of 98.9 % (95 % CI: 0.97–1.00). In both cases a single subsegmental pulmonary embolism was detected by gold standard reading. The CAD analysis performed during the study identified the embolism defined by gold standard reading as a finding in only one of the two cases. In the second case, the embolism defined by gold standard reading was not identified by the CAD software. Compared to the gold standard reading, no finding from routine reading at the emergency department was classified as false positive.

Duration of CAD analysis

In a sample of 50 datasets, the time from opening of the dataset to display of the CAD findings by the PE-CAD of the syngo.via software was 92.3 ± 15.5 seconds. The time needed for to evaluate the CAD findings varied depending on the number of CAD findings from 0 seconds in the case of no findings to 92 seconds in the case of the maximum of 25 CAD findings. On average, the evaluation of the CAD findings took 11.76 ± 16.27 seconds.

Contrast-to-noise ratio

The CNR was between 2.8 and 91.8 with a mean of 24.5 ± 12.6. The calculated CNR values are shown in relation to the number of datasets for the study population in ▶ **Fig. 5**. The Spearman's correlation between the CNR and the number of false-positive CAD findings (r_s = 0.085) did not show any correlation between the variables.

Discussion

The increasing availability and technical improvement of CT have resulted in increased use of CTPA in the diagnosis of pulmonary embolism [4]. Peripheral segmental and subsegmental Pes that can be overlooked particularly in the case of an isolated occurrence present a challenge here [6]. The relevance of an isolated subsegmental PE (ssPE) is a topic of controversy in the literature. While some studies have described overdiagnosis with subsequent overtreatment of ssPEs, the authors of other studies conclude that the clinical outcome of patients with subsegmental Pes and those with more proximal Pes is comparable [10–12]. Due to the unclear data, the valid guidelines of the European Society for Cardiology (ESC) recommend that treatment decisions regarding subsegmental Pes be made based on the patient's individual risk factors for PE and bleeding complication [13]. Thus, the detection of subsegmental Pes on CTPA continues to be highly relevant. Based on studies, CAD systems can support the analysis of CTPA datasets. However, the studies to date have evaluated small patient groups or CAD prototypes [6, 14–16].

    To our knowledge, this is the first study examining a commercial CAD system on the basis of a large patient population of a maximum care hospital in Germany and comparing CAD analysis with the findings of both emergency radiology and expert reading. In our study 1229 CTPA datasets were visually analyzed by two radiologists, which is referred to as gold standard reading. 504 embolisms were detected in 182 datasets and defined as the

▶ **Table 1** Overview of underlying cause of false-positive CAD findings as absolute value and as percentage. CAD - computer-aided detection.

cause for false-positive CAD findings	number of false-positive CAD findings
soft tissue	1876 (61 %)
pulmonary vein	742 (24.1 %)
pneumonia	357 (11.6 %)
atelectasis	59 (1.9 %)
contrast inflow	29 (0.9 %)
drainage	9 (0.3 %)
intrapulmonary calcification	1 (0.03 %)

▶ **Table 2** Fourfold table of gold standard reports and CAD findings on patient level. Gold std. - Gold standard; PE – pulmonary embolism; CAD - computer-aided detection..

	Gold std. PE detected	Gold std. No PE detected	total
at least 1 CAD finding	169 (92.9 %)	821 (78.4 %)	990
no CAD finding	13 (7.1 %)	226 (21.6 %)	239
total	182	1047	1229

gold standard for the evaluation of the CAD system. The PE-CAD - detected 258 embolisms corresponding to a sensitivity of 51.2 % on the embolism level. The sensitivity for the detection of segmental embolisms (53.9 %) was highest. The CAD software showed at least one true-positive finding in 118 of 182 patients with PE, corresponding to a sensitivity of 64.8 % on the patient level. These values are less than the ones described in the literature (sensitivity between 55.4 % and 76 % on the embolism level and between 83 % and 100 % on the patient level in recent studies) [14–19]. In our study, all routine findings from the emergency department were compared to the gold standard reading. Deviations could only be detected in two cases. In each case, one subsegmental embolism that was not documented in the clinical finding was identified by the gold standard reading. The CAD software identified the embolism in only one of the two cases. In the second case an additional CAD analysis also would not have been able to prevent the false-negative finding in the clinical routine.

    In patients with detected PE, at least one finding was identified with higher probability. On average, CAD analysis yielded 2.7 ± 2.61 findings in total and 2.5 ± 2.54 false-positive findings per patient. These values are significantly less than the values of other PE-CAD studies that range up to a maximum mean of 14.4 false-positive findings per patient [6, 16–18] and are slightly higher than the studies of Tajbakhsh et al. and Lahiji et al. with a mean
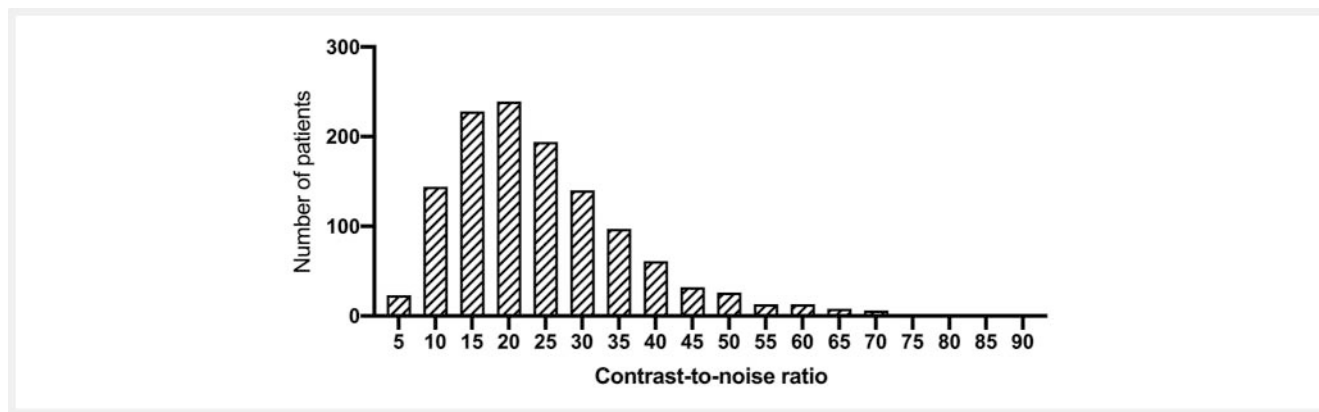
▶ **Fig. 5** Graph of the frequency distribution of contrast-to-noise ratio values.

of only 2 and 2.1 false-positive findings per dataset, respectively [14, 15]. The patient population alone cannot explain the significantly lower sensitivity in our study since Wittenberg et al. also analyzed a patient population of a university emergency department and reported a sensitivity of 95% on the patient level [16]. Moreover, an increased number of false-positive findings would be expected in a population comprised of patients in intensive care and the emergency department since the symptoms can be explained in a number of cases by other causes, e. g., pneumonia or atelectasis. It can thus be assumed that the low average rate of false-positive findings in our study in comparison to other studies is associated with a high rate of false-negative findings. In agreement with our results, mediastinal and hilar soft tissues as well as pulmonary veins are given as the most common causes of false-positive CAD findings [6, 15, 16].

In contrast to other studies, we were not able to show a significant correlation between image quality and the number of false-positive findings. However, it must be taken into consideration that the image quality was evaluated in the present study exclusively quantitatively on the basis of the CNR and not semiquantitatively on the basis of a Likert scale [6, 17, 20].

When using additional analysis software, the resulting time requirement is a relevant quality criterion. This is particularly true in the case of reading in an emergency setting. Therefore, we analyzed the additional time needed for the PE-CAD in a sample. The software showed the findings after an average of 92.3 ± 15.5 seconds. On average, an additional 11.76 ± 16.27 seconds were needed for the evaluation of the CAD findings. Therefore, dataset import and analysis by the CAD software require significantly more time than the evaluation of findings by a radiologist. Compared to Wittenberg et al. our CAD requires on average three times as long for the analysis [16]. However, since this process takes place at the same time as visual analysis in clinical application, no additional time is required. Contrary to general perception, evaluation of CAD findings only takes a few seconds, on average 11.76 ± 16.27 seconds, in the majority of cases. In a few cases, it took up to 92 seconds to evaluate the maximum number of 25 CAD findings.

The developments of recent years have illustrated that publicly available datasets are needed to be able to compare PE-CAD sys-

tems to one another [21]. In addition, studies have shown that there is still innovation potential. Therefore, monoenergetic reconstructions of dual energy datasets were used for CAD analysis, resulting in increased sensitivity [22, 23]. Three-dimensional segmentation of the pulmonary arterial circulation was also optimized [24, 25]. These developments indicate that with further optimization the PE-CAD can become part of our clinical routine.

The monocentric retrospective study design is a limitation of our study. Moreover, 601 datasets acquired as a dual rule out protocol for simultaneously ruling out a PE and an aortic dissection had to be excluded. These datasets could not be imported into the PE-CAD. Therefore, a selection bias is possible.

Conclusion

Evaluation of the commercial PE-CAD system based on a large patient population in an emergency setting showed a need to optimize the software, particularly in relation to false-negative findings. The study highlights the role of the system as supporting analysis software that cannot (yet) replace visual assessment by radiologists.

CLINICAL RELEVANCE OF THE STUDY

- Since PE is a common disease with potentially high morbidity and mortality, CTPA is frequently performed in the emergency setting.
- In particular, isolated subsegmental Pes can be overlooked.
- PE-CAD systems can support radiologists during analysis without a significant increase in the time requirement.
- When using these systems, it must be taken into consideration that CAD analysis generates both false-negative and false-positive findings and cannot replace visual evaluation.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Calder KK, Herbert M, Henderson SO. The mortality of untreated pulmonary embolism in emergency department patients. Ann Emerg Med 2005; 45: 302–310. doi:10.1016/j.annemergmed.2004.10.001

[2] Wendelboe AM, Raskob GE. Global Burden of Thrombosis: Epidemiologic Aspects. Circ Res 2016; 118: 1340–1347. doi:10.1161/circresaha.115.306841

[3] Diehm C, Noppeney T, Nüllen H. Epidemiologie der venösen Thromboembolie. Gefässchirurgie 2012; 17: 275–279. doi:10.1007/s00772-011-0929-8

[4] Sharma S, Lucas CD. Increasing use of CTPA for the investigation of suspected pulmonary embolism. Postgrad Med 2017; 129: 193–197. doi:10.1080/00325481.2017.1281084

[5] Schissler A, Rozenshtein A, Kulon M et al. CT Pulmonary Angiography: Increasingly Diagnosing Less Severe Pulmonary Emboli. PloS one 2013; 8: e65669 doi:10.1371/journal.pone.0065669

[6] Kligerman SJ, Lahiji K, Galvin JR et al. Missed pulmonary emboli on CT angiography: assessment with pulmonary embolism-computer-aided detection. Am J Roentgenol 2014; 202: 65–73. doi:10.2214/ajr.13.11049

[7] Hawley PC, Hawley MP. Difficulties in diagnosing pulmonary embolism in the obese patient: a literature review. Vasc Med 2011; 16: 444–451. doi:10.1177/1358863x11422571

[8] Buhmann S, Herzog P, Liang J et al. Clinical evaluation of a computer-aided diagnosis (CAD) prototype for the detection of pulmonary embolism. Acad Radiol 2007; 14: 651–658. doi:10.1016/j.acra.2007.02.007

[9] Moore AJE, Wachsmann J, Chamarthy MR et al. Imaging of acute pulmonary embolism: an update. Cardiovasc Diagn Ther 2018; 8: 225–243. doi:10.21037/cdt.2017.12.01

[10] Eyer BA, Goodman LR, Washington L. Clinicians' response to radiologists' reports of isolated subsegmental pulmonary embolism or inconclusive interpretation of pulmonary embolism using MDCT. Am J Roentgenol 2005; 184: 623–628. doi:10.2214/ajr.184.2.01840623

[11] Wiener RS, Schwartz LM, Woloshin S. Time trends in pulmonary embolism in the United States: evidence of overdiagnosis. Arch Intern Med 2011; 171: 831–837. doi:10.1001/archinternmed.2011.178

[12] den Exter PL, van Es J, Klok FA et al. Risk profile and clinical outcome of symptomatic subsegmental acute pulmonary embolism. Blood 2013; 122: 1144–1149; quiz 1329. doi:10.1182/blood-2013-04-497545

[13] Konstantinides SV, Meyer G, Becattini C et al. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS): The Task Force for the diagnosis and management of acute pulmonary embolism of the European Society of Cardiology (ESC). European Heart Journal 2019; 41: 543–603. doi:10.1093/eurheartj/ehz405

[14] Tajbakhsh N, Gotway MB, Liang J. Computer-Aided Pulmonary Embolism Detection Using a Novel Vessel-Aligned Multi-planar Image Representation and Convolutional Neural Networks. Cham: Springer International Publishing; 2015: 62–69

[15] Lahiji K, Kligerman S, Jeudy J et al. Improved accuracy of pulmonary embolism computer-aided detection using iterative reconstruction compared with filtered back projection. Am J Roentgenol 2014; 203: 763–771. doi:10.2214/ajr.13.11838

[16] Wittenberg R, Peters JF, Sonnemans JJ et al. Computer-assisted detection of pulmonary embolism: evaluation of pulmonary CT angiograms performed in an on-call setting. Eur Radiol 2010; 20: 801–806. doi:10.1007/s00330-009-1628-7

[17] Wittenberg R, Peters JF, Weber M et al. Stand-alone performance of a computer-assisted detection prototype for detection of acute pulmonary embolism: a multi-institutional comparison. Br J Radiol 2012; 85: 758–764. doi:10.1259/bjr/26769569

[18] Özkan H, Osman O, Şahin S et al. A novel method for pulmonary embolism detection in CTA images. Comput Methods Programs Biomed 2014; 113: 757–766. doi:10.1016/j.cmpb.2013.12.014

[19] Krissak R, Henzler T, Reichert M et al. Enhanced visualization of lung vessels for diagnosis of pulmonary embolism using dual energy CT angiography. Invest Radiol 2010; 45: 341–346. doi:10.1097/RLI.0-b013e3181dfda37

[20] Wittenberg R, Peters JF, Sonnemans JJ et al. Impact of Image Quality on the Performance of Computer-Aided Detection of Pulmonary Embolism. American Journal of Roentgenology 2011; 196: 95–101. doi:10.2214/Am J Roentgenol.09.4165

[21] Masoudi M, Pourreza HR, Saadatmand-Tarzjan M et al. A new dataset of computed-tomography angiography images for computer-aided detection of pulmonary embolism. Sci Data 2018; 5: 180180 doi:10.1038/sdata.2018.180

[22] Ma G, Dou Y, Dang S et al. Influence of Monoenergetic Images at Different Energy Levels in Dual-Energy Spectral CT on the Accuracy of Computer-Aided Detection for Pulmonary Embolism. Acad Radiol 2019; 26: 967–973. doi:10.1016/j.acra.2018.09.007

[23] Kröger JR, Hickethier T, Pahn G et al. Influence of spectral detector CT based monoenergetic images on the computer-aided detection of pulmonary artery embolism. Eur J Radiol 2017; 95: 242–248. doi:10.1016/j.ejrad.2017.08.034

[24] Tajbakhsh N, Shin JY, Gotway MB et al. Computer-aided detection and visualization of pulmonary embolism using a novel, compact, and discriminative image representation. Med Image Anal 2019; 58: 101541 doi:10.1016/j.media.2019.101541

[25] Zhang C, Sun M, Wei Y et al. Automatic segmentation of arterial tree from 3D computed tomographic pulmonary angiography (CTPA) scans. Comput Assist Surg (Abingdon) 2019; 24: 79–86. doi:10.1080/24699322.2019.1649077