Computer-Assisted Detection of Acute Pulmonary Embolism

Rianne Wittenberg

Computer-Assisted Detection of Acute Pulmonary Embolism
Thesis, University of Utrecht, The Netherlands

Cover : Karin van Rijnbach and Rianne Wittenberg
Lay-out : Karin van Rijnbach
Printed by : PrintSupport4U
ISBN : 978-90-393-5736-1

Computer-Assisted Detection of Acute Pulmonary Embolism

Automatische Detectie van Acute Longembolieën

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op donderdag 22 maart 2012 des ochtends te 10.30 uur

door

Rianne Wittenberg
geboren op 1 november 1982 te Hilversum

Promotoren: Prof. dr. W.M. Prokop
Prof. dr. J.S. Laméris

Co-promotor: Dr. C.M. Schaefer-Prokop

Publication of this thesis was financially supported by:
Philips Healthcare, Academic Medical Centre Amsterdam and Röntgen Stichting Utrecht

Contents

Chapter 1

General introduction and thesis outline

Introduction

Pulmonary embolism

Pulmonary embolism (PE) is a very common, potentially life threatening disease presenting with non-specific clinical symptoms like fever, coughing, dyspnea and chest pain. In western populations, the incidence varies between 1.22 and 1.83 per 1.000 among adults [1] and these numbers are likely underestimated because most people die of other underlying diseases. The prevalence of PE detected at autopsy among hospitalised patients amounts to 15-26% [2,3].

Pulmonary arterial obstruction may cause an increase in the right ventricular afterload and may therefore lead to dilatation, dysfunction and ischemia of the right ventricle resulting in death [4]. Mortality rates of up to 25% are reported if PE is not treated in time [5-7]. Several risk factors for the development of PE are known (Table 1). Knowledge of these risk factors could help physicians to estimate the disease probability.

Table 1. Risk factors for PE

Acute medical illness
Thrombophilia
Cancer
Major surgery
Trauma
Immobility
Age
Obesitas
Cigarette smoking
Oral contraceptives
Pregnancy and postpartum period
Hormone replacement therapy
Central venous catheterisation

The first line treatment for PE is anticoagulation therapy based on a dichotomous “yes” or “no” decision for the presence of PE irrespective of the number, size or distribution of thrombi. Only a very small subgroup of patients with obstructing central emboli will undergo immediate surgical thrombectomy or interventional procedures for removal of central clots. Unfortunately, because of the non-specific clinical symptoms, PE is clinically difficult to diagnose and

medical imaging is necessary. Over the last decades, several diagnostic methods have been used for the detection of PE.

Diagnostic methods

A planar ventilation/perfusion (V/Q) scintigram is a safe and widely available technique for the diagnosis of PE. Before the introduction of spiral CT, VQ scintigraphy had been the first-line test for diagnosing PE for more than 30 years. Normal results on a V/Q scan securely rule out PE and a high probability lung scan is diagnostic for PE with a probability of more than 90%. However, planar V/Q scintigraphy provides limited alternative diagnostic information and because 60-70% of the V/Q scans are non-diagnostic, additional tests were often necessary [8,9]. If the diagnosis was still unclear, patients would undergo pulmonary angiography, which was the golden standard for the diagnosis of PE. This is an invasive procedure and although the number of complications are low, they are severe [10]. When at the end of the 1990s Diffin et al. and Stein at al. [11,12] also showed that the diagnostic accuracy of pulmonary angiography was rather low for the diagnosis of peripheral emboli, the need for a new, accurate and non-invasive technique increased.

Single photon emission computed tomography (SPECT) is considered as the advanced successor of the planar V/Q lung scan. In addition to the information on perfusion, SPECT offers the advantage of cross sectional imaging with increased spatial resolution and decreased effects of super projection. According to the literature, the number of non-diagnostic tests decreases to less than 5% with the use of V/Q SPECT. However, in the literature, variable methods for obtaining SPECT images are reported, different reference standards are used and the criteria for interpretation of low and intermediate probability scans differ substantially [13,14]. As compared to CT, SPECT is less available and in many institutions not in use on a 24/7 basis. Besides, CT offers a diagnosis for PE within seconds and options alternative diagnosis. Therefore, computed tomography angiography became the first method of choice for the diagnosis of PE.

Computed tomography angiography

In the 1970s, Sir Godfrey Hounsfield (Hayes, United Kingdom) developed the first CT scanner using Cormack's theoretical calculations for X-ray absorption in tissue. For their independent efforts, Cormack and Hounsfield shared the 1979 Nobel Prize in Physiology or Medicine. In 1989, the first spiral computed tomography pulmonary angiography (SCTPA) was demonstrated, followed by the multidetector computed tomography (MDCT) in 1992. The first prospective study comparing SCTPA with pulmonary angiography was published by Remy-

Jardin et al. in 1992 [15]. They concluded that spiral volumetric CT could reliable depict thromboemboli in second- to fourth-division pulmonary vessels.

Over the last two decades, CT technology has dramatically improved. The sensitivity and specificity for the diagnosis of PE vary between 83% and 100% and 89% and 97% respectively [16,17]. With the first MDCT scanners, spatial resolution was limited by the trade off between scan range on one side and spatial resolution on the other and could be achieved within a single breath-hold. Nowadays, with the new generation MDCT scanners with at least 64-detector rows, the spatial resolution has significantly improved and allows us to depict PE even at the fifth or sixth level of the arteries [9]. This has lead to the general acceptance of the MDCT as first choice modality for the detection of PE. However, there are also some disadvantages of the MDCT such as contraindications to iodinated contrast and the moderate amount of irradiation. Furthermore, with the improved spatial resolution offered by MDCT, the number of slices to be evaluated by the radiologists continuously increased. Not only the growing numbers of examinations and slices, but also the increasing computer power have lead to the introduction of computer-assisted diagnosis designed to support the radiologist in detecting and/or diagnosing pathology.

Computed-assisted detection

Driven by the increasing potential, computer-assisted detection (CAD) or computer-assisted diagnosis became an important research subject in diagnostic radiology. Computer-assisted detection systems allow for the automatic identification of certain lesions, providing another opinion to the radiologists' judgement. These systems, for example for lung nodules, are able not only to identify them but also to characterise them by determining a likelihood of malignancy or benignity. So far, CAD for various radiologic applications has been investigated including mammography, thoracic CT and virtual colonoscopy [18-20].

Computer-assisted detection for the diagnosis of PE is relatively new. The first study on that subject was published in 2002 by Masutani et al [21]: for a small group of 19 selected CTPA examinations with good image quality, they reported a sensitivity on a per embolus basis of 85% with 2.6 false positive findings per examination. Since then, various CAD algorithms for the detection of PE have been tested by different vendors and image analysis groups. Besides one very early version, none of them have been FDA approved until today or have been tested in a clinical environment. Before newly developed software can be implemented in clinical practice the performance of this software needs to be investigated.

The performance of CAD software can be tested in different ways:
1) The standalone performance refers to the performance of the technique in

a defined set of clinical data without interference of human observers (in our setting radiologists or radiologists-in-training).
2) The performance of the algorithm can be tested within observer studies, meaning that the impact of the CAD analysis on the performance of human observers is evaluated. A CAD algorithm can be used in various ways. It can be used as a "second reader" meaning that the CAD output is only available for the reader after he/she has assessed the examination primarily unassisted and uses the results of CAD only to refine his/her judgment. In the following, it is debatable whether the CAD output can only be used to add lesions but does not allow to remove lesions originally seen by the readers. In a concurrent reading setting, CAD findings are immediately displayed to the reader already during the first visual assessment of the examination.

Most studies so far mainly focused on the standalone performance of CAD algorithms [21-26]. In the beginning, most performance parameters were calculated on a per lesion basis. Because anticoagulation therapy is clinically based on a "yes" or "no" decision for the diagnosis of PE, it seems to be more relevant from a clinical point of view to assess performance parameters on a per patient basis. However, not only the standalone performance but also the effect of CAD on readers' decisions needs to be investigated. Only four studies have been published with CAD as a second reader using rather small patient groups or using selected groups of scans with only good quality or with only positive patients [27-30]. Furthermore, using CAD as a concurrent reader for the detection of PE has never been investigated before.

Thesis outline

The purpose of this thesis was to assess the standalone performance of a CAD prototype (Philips Healthcare, Best, The Netherlands) for the detection of acute pulmonary embolism and the impact of this system on readers' performance.

Part 1: Update on detecting pulmonary embolism using CT angiography
Since its introduction in 1992, CTPA has become the reference standard in patients suspected of having pulmonary embolism (PE). The development of the MDCT has lead to faster scanning protocols and major advances in resolution and image quality. In **chapter 2**, an introduction into recent technical advances and its practical implications is given for CTPA for the diagnosis of PE.

Part 2: Standalone performance of a computer-assisted detection prototype for the detection of acute pulmonary embolism
The first study described in **chapter 3** assessed the standalone performance of the CAD prototype to provide an estimation of its potential to support readers for diagnosing and excluding PE. We expected that the CAD performance would depend on the image quality of the scan. In **chapter 4**, we therefore assessed the relation between the standalone performance and the image quality of the scan. In clinical routine, the image quality of a scan strongly depends on CT protocol parameters, CT vendor and patient instructions. Institutions use their own implemented CT protocols and different scanner types. To examine whether the results described in chapters 3 and 4 were also valid in other hospitals, we compared the standalone performance of this CAD prototype in CTPA examinations obtained in 3 different institutions and described the results in **chapter 5.**

Part 3: Impact of a computer-assisted detection prototype used as a second or a concurrent reader on readers with varying experience
Whether CAD can improve the performance of (in)experienced readers will not only depend on the standalone performance, but also on readers' experience, their confidence in the system and their ability to differentiate true from false positive CAD findings. To show the impact of this CAD prototype on readers' decisions, 4 radiology residents and 2 radiologists of varying experience tested the CAD prototype as a second reader using 209 consecutive CTPA scans. These results are described in **chapter 6**. CAD software is initially developed to use as a "second reader", meaning that CAD candidates are made available to the readers only after having made a complete interpretation of the scan unassisted by the software. This reading method, however, carries the risk

that readers may do the first analysis less consciously and will inevitably lead to an increase of reading time. An alternative reading method is to have CAD candidates already available during the initial analysis, thus using CAD as a concurrent reader. In **chapter 7**, we therefore investigated the impact of CAD as a concurrent reader on 3 radiology residents and 3 radiologists, using 196 consecutive CTPA examinations.

The results of this thesis and its implications for clinical routine are discussed in **chapter 8**.

Reference List

1. Goldhaber SZ. Pulmonary embolism. Lancet 2004; 363(9417):1295-305.
2. Lindblad B, Eriksson A, Bergqvist D. Autopsy-verified pulmonary embolism in a surgical department: analysis of the period from 1951 to 1988. Br J Surg 1991; 78(7):849-52.
3. Stein PD, Henry JW. Prevalence of acute pulmonary embolism among patients in a general hospital and at autopsy. Chest 1995; 108(4):978-81.
4. Tapson VF. Acute pulmonary embolism. N Engl J Med 2008; 358(10):1037-52.
5. Cushman M, Tsai AW, White RH et al. Deep vein thrombosis and pulmonary embolism in two cohorts: the longitudinal investigation of thromboembolism etiology. Am J Med 2004; 117(1):19-25.
6. Goldhaber SZ, Visani L, De RM. Acute pulmonary embolism: clinical outcomes in the International Cooperative Pulmonary Embolism Registry (ICOPER). Lancet 1999; 353(9162):1386-9.
7. Heit JA, Silverstein MD, Mohr DN et al. Predictors of survival after deep vein thrombosis and pulmonary embolism: a population-based, cohort study. Arch Intern Med 1999; 159(5):445-53.
8. Value of the ventilation/perfusion scan in acute pulmonary embolism. Results of the prospective investigation of pulmonary embolism diagnosis (PIOPED). The PIOPED Investigators. JAMA 1990; 263(20):2753-9.
9. Remy-Jardin M, Pistolesi M, Goodman LR et al. Management of suspected acute pulmonary embolism in the era of CT angiography: a statement from the Fleischner Society. Radiology 2007; 245(2):315-29.
10. Stein PD, Athanasoulis C, Alavi A et al. Complications and validity of pulmonary angiography in acute pulmonary embolism. Circulation 1992; 85(2):462-8.
11. Diffin DC, Leyendecker JR, Johnson SP et al. Effect of anatomic distribution of pulmonary emboli on interobserver agreement in the interpretation of pulmonary angiography. AJR Am J Roentgenol 1998; 171(4):1085-9.
12. Stein PD, Henry JW, Gottschalk A. Reassessment of pulmonary angiography for the diagnosis of pulmonary embolism: relation of interpreter agreement to the order of the involved pulmonary arterial branch. Radiology 1999; 210(3):689-91.
13. Duc-Pennec LE, LE Roux PY, Cornily JC et al. Diagnostic accuracy of single photo emission tomography ventilation perfusion (SPECT V/Q) lung scan in the diagnosis of pulmonary embolism. Chest 2011 [Epub ahead of print].
14. Stein PD, Freeman LM, Sostman HD et al. SPECT in acute pulmonary embolism. J Nucl Med 200950(12):1999-2007.
15. Remy-Jardin M, Remy J, Wattinne L et al. Central pulmonary thromboembolism: diagnosis with spiral volumetric CT with the single-breath-hold technique--comparison with pulmonary angiography. Radiology 1992; 185(2):381-7.
16. Stein PD, Fowler SE, Goodman LR et al. Multidetector computed tomography for acute pulmonary embolism. N Engl J Med 2006; 354(22):2317-27.

17. Winer-Muram HT, Rydberg J, Johnson MS et al. Suspected acute pulmonary embolism: evaluation with multi-detector row CT versus digital subtraction pulmonary arteriography. Radiology 2004; 233(3):806-15.
18. Chan HP, Hadjiiski L, Zhou C et al. Computer-aided diagnosis of lung cancer and pulmonary embolism in computed tomography-a review. Acad Radiol 2008; 15(5):535-55.
19. Malich A, Fischer DR, Bottcher J. CAD for mammography: the technique, results, current role and further developments. Eur Radiol 2006; 16(7):1449-60.
20. Petrick N, Haider M, Summers RM et al. CT colonography with computer-aided detection as a second reader: observer performance study. Radiology 2008; 246(1):148-56.
21. Masutani Y, MacMahon H, Doi K. Computerized detection of pulmonary embolism in spiral CT angiography based on volumetric image analysis. IEEE Trans Med Imaging 2002; 21(12):1517-23.
22. 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(6):651-8.
23. Dewailly M, Remy-Jardin M, Duhamel A et al. Computer-aided detection of acute pulmonary embolism with 64-slice multi-detector row computed tomography: impact of the scanning conditions and overall image quality in the detection of peripheral clots. J Comput Assist Tomogr 2010; 34(1):23-30.
24. Maizlin ZV, Vos PM, Godoy MC et al. Computer-aided detection of pulmonary embolism on CT angiography: initial experience. J Thorac Imaging 2007; 22(4):324-9.
25. Schoepf UJ, Schneider AC, Das M et al. Pulmonary embolism: computer-aided detection at multidetector row spiral computed tomography. J Thorac Imaging 2007; 22(4):319-23.
26. Zhou C, Chan HP, Patel S et al. Preliminary investigation of computer-aided detection of pulmonary embolism in three-dimensional computed tomography pulmonary angiography images. Acad Radiol 2005; 12(6):782-92.
27. Blackmon KN, Florin C, Bogoni L et al. Computer-aided detection of pulmonary embolism at CT pulmonary angiography: can it improve performance of inexperienced readers? Eur Radiol 2011; 21(6):1214-23.
28. Das M, Muhlenbruch G, Helm A et al. Computer-aided detection of pulmonary embolism: Influence on radiologists‘ detection performance with respect to vessel segments. Eur Radiol 2008; 18:1350-5.
29. Engelke C, Schmidt S, Bakai A et al. Computer-assisted detection of pulmonary embolism: performance evaluation in consensus with experienced and inexperienced chest radiologists. Eur Radiol 2008; 18(2):298-307.
30. Walsham AC, Roberts HC, Kashani HM et al. The use of computer-aided detection for the assessment of pulmonary arterial filling defects at computed tomographic angiography. J Comput Assist Tomogr 2008; 32(6):913-8.

Part 1

Update on detecting pulmonary embolism using CT angiography

Chapter 2

Hartmann IJ
Wittenberg R
Schaefer-Prokop CM

Eur J Radiol. 2010 Apr;74(1):40-9.

Imaging of acute pulmonary embolism using multidetector CT angiography: an update on imaging technique and interpretation

Abstract

Computed tomography angiography (CTA) of the pulmonary arteries has become the main diagnostic test for the evaluation of pulmonary embolism (PE). This is not merely due to the good availability, low cost and minimal invasiveness of this technique, but mainly because of the introduction of multidetector CT techniques resulting in significant improvement in resolution, speed and image quality. This continuous gain in image acquisition speed went along with the introduction of new techniques of image acquisition, such as the dual-source CT scanning and novel concepts of image interpretation beyond morphological findings, including the definition of the resulting perfusion defects and assessment of the cardiopulmonary circulation as a functional unit. This article will focus on technical and practical aspects to optimise CTPA examinations with modern multidetector CT scanners, discusses aspects to be considered in specific patient groups (e.g. during pregnancy, young patients) and outlines new advents such as dual-source lung perfusion and automatic detection of PE.

Introduction

Since its introduction in 1992, computed tomography angiography (CTA) of the pulmonary arteries has become the main diagnostic test for the evaluation of pulmonary embolism (PE) [1]. With the advent of multidetector scanning, CTA has gained substantially in image acquisition speed and spatial resolution, which changed its diagnostic yield in many respects. With a single slice CT scanner, a scan volume of 16 cm could be scanned with 5 mm collimation in 32 s, whereas with the newest generation 64-row detector scanners, the complete chest of 30 cm is covered in less than 5 s with sub-millimeter collimation. This continuous gain in image acquisition speed went along with the introduction of new techniques of image acquisition, such as the dual-source CT scanning and novel concepts of image interpretation beyond morphological findings, including the definition of the resulting perfusion defects and assessment of the cardiopulmonary circulation as a functional unit. With the development of these fast multidetector scanning techniques, the number of non-diagnostic examinations substantially decreased (up to 10% for single detector CT versus up to 6% for multidetector CT) [2,3]. However, a qualitatively suboptimal examination resulting in an ambiguous interpretation is still the most important drawback of CT for the detection of acute PE and optimisation of acquisition techniques, including the contrast injection protocol remains important also with the most modern scanners.

This article will focus on technical and practical aspects to optimise CTPA examinations with modern multidetector CT scanners, discusses aspects to be considered in specific patient groups (e.g. during pregnancy, young patients) and outlines new advents such as dual-source lung perfusion and automatic detection of PE.

Acquisition techniques

The choice of the optimal imaging protocol depends on the equipment at hand, because 4 to 256 or even 320-row scanners are currently in practical use. While with 4 and 8 detector row scanners, an optimal compromise between scan duration, breath-hold capability and slice collimation has to be sought for, 16 up to 64 detector row scanners allow for examining the whole chest in less than 5 s with a collimation of at least 1 mm or even lower and faster scanners examine the chest in less than a second. Thus, even in a dyspneic patient, the need for a long breath-hold capability is not an issue anymore. Depending on the scanner used, the thinnest collimation for multidetector spiral CT acquisitions ranges

from 0.5 to 1 mm, enabling one to provide the most detailed display of the pulmonary arteries down to the subsegmental level routinely and to improve the detection rate of segmental and subsegmental PE due to reduced volume averaging [4]. Other positive effects are a significant decrease of indeterminate scan results and improved interobserver agreement [5,6]. However, these ultra short scanning times make optimisation of the contrast application and patient instruction even more critical and important. Even small movements, a short Valsalva or a minor change of delay can lead to a complete inadequate examination.

Patient instruction

It is recommended to perform data acquisition during apnea, preferably at total lung capacity. In cooperative patients without dyspnea the scan is performed during an inspiratory breath hold. With longer scanning times (>15 s), a short hyperventilation before scanning acquisition may be helpful. Careful patient instruction by the technologists is warranted and performance of a trial breath hold for the required scan duration may be helpful [7]. By watching the surface of the abdomen, adequate suspension of respiration can be assessed. Patients

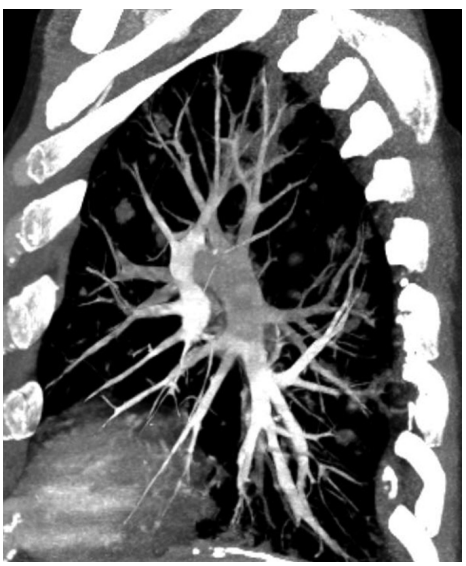

Figure 1. CT scan performed in caudo-cranial direction. The sagittal maximum intensity projections demonstrates good vascular enhancement of the pulmonary arteries in the lower parts of the lung but no enhancement of the pulmonary arteries in the upper parts of the lung due to a Valsalva manoeuvre.

require usually 4 s between the breath hold command and the actual scan before they have fully suspended diaphragmatic motion. Substantial movement artefacts may occur if the start delay is too short or patients are improperly instructed. In addition, inspiration immediately prior to the imaging acquisition may lead to a transient interruption of the contrast column in the pulmonary arteries. This is the result from a variable inflow of unopacified blood from the inferior vena cava as a normal response to the negative intrathoracic pressure [8]. In contrary, Valsalva manoeuvres lead to diminished inflow of the contrast medium column as a result of positive intrathoracic pressure (Figure 1). Both may lead to inadequate and inhomogeneous intravascular contrast that becomes even more critical with the scanners that need less than 5 s to cover the chest. Therefore, technologists need to be instructed not to start the scans too early and patients to take a deep breath and hold their breath without performing a Valsalva manoeuvre. Some patients tend to 'gasp for air' internally against the closed glottis at the end of the scan acquisition. This may result in motion artefacts due to the involuntary diaphragmatic movements and can be avoided by careful instruction of the patient. In heavily dyspneic patients, shallow breathing is preferred over forced breath holding to avoid heavy and more uncontrollable movement artefacts.

Scan range and direction

Routinely, CTA of the pulmonary arteries normally includes the complete chest, including long apex and costophrenic angles. Data so far indicates that limitation of the scan range, e g. from the dome of the lower hemidiaphragm to 2–4 cm above the aortic arch will not lead to a decrease of detected PE. Limitation of the scan length in that way to some 16–24 cm, will lead to a reduction of radiation exposure and scan duration by 20–45% compared to a full chest scan of 30 cm, which makes it a valuable approach for patients with increased radiation sensitivity. Detection and extent of accompanying disease may be partially missed though, thus this approach is not recommended in patients with comorbidity. With single-slice CTA, most institutions preferred a caudocranial scanning direction because breathing artefacts during the final phase of data acquisition were less disturbing in the upper portions of the lung and to avoid beam hardening artefacts through inflow of high contrast. With shorter scanning times and use of a saline chaser, these aspects have become less of an issue. Today, both caudo-cranial and cranio-caudal scanning directions are in use, dependant on local preference.

Scanning parameters

Scanning parameters depend on the available scanner technology and the breath hold capacity of the patient. Table 1 summarises acquisition protocols for different scanner types. Only for scanners with less than 16 detector rows, it is necessary to adapt slice collimation and scanning time to the patient's breath hold capability. With the faster scanners, scanning time has become so short that a sacrifice of spatial resolution for scanning time is not an issue anymore. Depending on the slice collimation during acquisition, slice reconstruction width varies between 0.9 and 1.5 mm. In adipose patients, a smoothing reconstruction algorithm and thicker slices between 1.5 or 2 mm may be advantageous to increase signal to noise. It is not recommended to base the diagnosis on slices thicker than 2 mm to ensure optimal evaluation of peripheral vascular structures.

Dual-source CT scanners can be used in two different ways for the detection of acute PE. Using the Flash technique, the chest can be scanned from the apex to the diaphragm with the thinnest collimation in less than 1 s. This will further reduce the risk for movement artefacts and increases the temporal resolution. Second, dual-source CT scanner offers the option to reconstruct "material specific", e.g. images in which the distribution of the iodium is used to produce lung perfusion images [9-12]. Reconstruction of the two data sets obtained at two different kilovoltages as fused images produces a "standard" CTA. Subtracting the lower kV from the higher kV images using dedicated processing software produces CT-perfusion colour-coded maps, resembling the distribution of iodine in the lung. To allow the contrast material to perfuse the complete lung, the delay time should not be too short and a saline flush is needed to limit streak artefacts from the superior vena cava. Although the perfusion technique seems promising for combining direct visualisation of the thrombus with the functional consequences, namely perfusion defects, the additional diagnostic value of this

Table 1. Scan parameters for the different types of scanners.

Scanner type	Collimation (mm)	Rotation time (s)	Scan duration (s, 24 cm)
4	1–1.25	0.5	≤20
4 (dyspnea)	2–2.5	0.8	≤10
16	0.625–1	0.37–0.42	≤8
64	0.5–0.625	0.35–0.42	≤4
128	0.6	0.30	<3
128 DS	0.6	0.28	<1
256	0.625	0.33	< 2
320 (64)	0.5	0.35	< 4

Table 2. Scan parameters for dual-source CT scanners.

	Dual-energy		**High speed (flash)**
	Tube A	Tube B	
kV	80–100	140	100-120
mAsref	125–180	125-180	170-200
Detector row	64	64	128
Collimation	0.6	0.6	0.3
Rotation time	0.28	0.28	0.28
Pitch	0.7	0.7	0.6
Scan time	24 cm ≤ 10 s	-	24 cm ≤ 1 s
Contrast injection (flow Iomeron 400)	75 ml (5 ml/s)	-	80-100 ml (5-6 ml/s)
NaCl (flow)	45 ml (5 ml/s)	-	-

technique still needs to be determined in future studies. Different dual-source protocols for the assessment of PE are provided in Table 2.

ECG-gating

Some scanners are capable of prospective ECG-gating. ECG-gating diminishes motion artefacts due to cardiac pulsation, especially in the lingula and left lower lobe vessels and may therefore improve image quality as compared to non-ECG-gated CTA. Whether ECG-gating actually increases diagnostic accuracy in a clinically relevant manner is doubtful and therefore seems not to be indicated on a routine basis [13]. More important is that ECG-gated CTA also allows for the assessment of right ventricular function that is likely to play a role for the outcome of the patient in association with the thrombus load. Last but not least, a triple or dual rule out protocol in patients with acute chest pain requires ECG triggering for sufficient evaluation of the coronary arteries [14].

Injection protocols

Injection parameters

Objective of the injection protocol is to ensure a constant and high degree of pulmonary arterial opacification during the entire image acquisition. This becomes more challenging the shorter the scan time is because Valsalva manoeuvres or inadequate delay times can destroy the whole scan. Injection of contrast medium

Table 3. Injection protocols for the different types of scanners.

Scan duration (s)	Scanner type	With saline (ml)/(ml/s)/(s)	Without saline ml
≤5	16, ≥64	70 + 40/5/8-10P*	80
10	4 (dyspnea), 16	80 + 40/5/5-8P*	100
20	4	100 + 40/4/5P*	120

* *P*, the time that should be added after the threshold in the pulmonary artery is reached using bolus triggering.

with high flow rates (4–6 ml/s) and of high-concentration (370–400 mg of iodine/ml) result in a high delivery of mg iodine per s, which enhances visualisation of small, more peripheral vessels and thus improves the overall sensitivity of the technique [15]. As opposed to organ imaging (e.g. the liver) in CTA of the pulmonary arteries the column of contrast can be substantially decreased with faster scanners. The duration of contrast injection should be approximately equal to the sum of the scan duration and the delay time. For very short scan durations, one has to increase the delay time to allow for adequate contrast build-up in the complete scan range and in the peripheral pulmonary arteries (Table 3). Mostly, a saline chaser of 30–60 ml is immediately injected after the contrast medium using the same flow rate to flush the accessory veins to decrease the risk for beam hardening artefacts, to prolong the length of the contrast plateau and to decrease the total volume of contrast medium needed [16].

With the faster scanners and the consequently short scanning time, an individualised bolus triggering becomes essential. This also allows for adequate adaptation to a non-expected circulatory slow down due to right heart failure, pulmonary hypertension or low cardiac output but also to a hypercirculatory status, e.g. in young patients or in pregnant women. After reaching the trigger threshold located in the right ventricle or the pulmonary trunk, an extra 5–8 s will have to be added to allow for an adequate build-up of intravascular contrast. The delay is longer and becomes more crucial the faster the scanner and the shorter the scan time is. As already pointed out, adequate patient instruction is mandatory.

Right cubital venous access is preferred for pulmonary CTA to avoid streak artefacts caused by the approximately horizontal course of the left brachiocephalic vein (Figure 2). In general, both arms are placed above the head. Some authors prefer to position the left arm above the head and the right arm parallel to the body when using right cubital venous access for the contrast injection. This technique reduces compression effects of the brachiocephalic vein when crossing the ribs. However, placing the arm parallel to the body may lead to increased

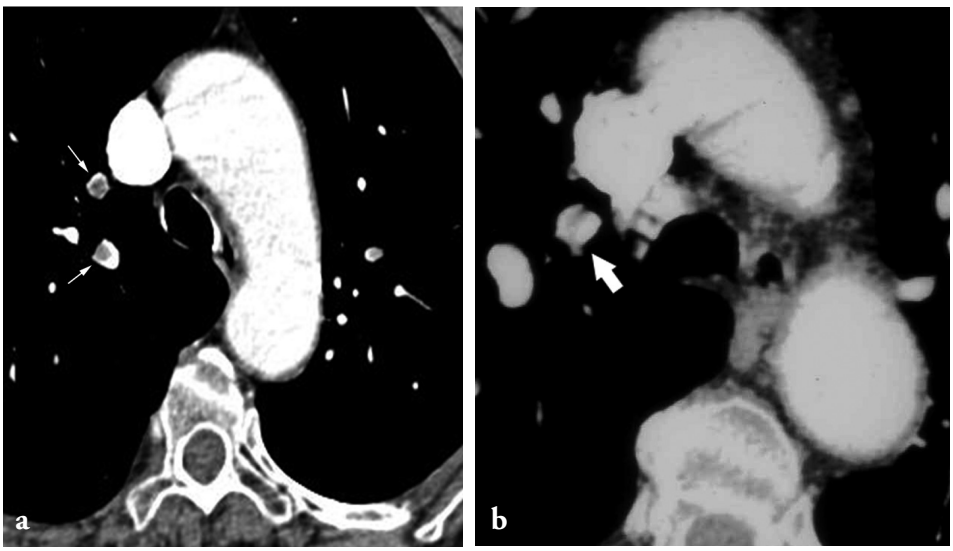

Figure 2. Bilateral filling defects in pulmonary arteries of the upper lobes due to pulmonary embolism (**a** - small arrows) and beam hardening artefacts (**b** - large arrow).

noise in the scanned volume if low-dose protocols are used.

The speed of data acquisition available with at least 16-row detector scanners allows for the use of gadolinium in selected patients with a suspicion for PE who have contraindications to iodinated contrast medium injection [17,18]. With the limited amount of gadolinium that can be administered for pulmonary CTA, namely 3–4 mmol/kg, adequate opacification of the pulmonary arteries can be obtained during the complete image acquisition, enabling confident detection of PE in even subsegmental pulmonary arteries.

Radiation dose considerations

In recent years, multiple efforts have been made by both the manufacturer and the user to minimise the dose required for CTA, without losing image quality that may impair diagnosis. The ALARA principle has become the driving force. More so since often times, young patients are examined and the vast majority of these examinations are done to rule out PE. The use of dose modulation software – offered by all manufacturers – is standard. It enables the tube current and hence the dose applied to be adjusted to the individual patient's geometry and patient's absorption during the data acquisition, i.e. it compensates for varying dose requirements between different body cross sections such as the shoulder regions and the air-filled chest. At a constant image quality, adaptive tube current modulation can reduce the dose by some 10–50%, depending on the body region (shoulders, central chest, diaphragm) [19,20]. In addition, the dose

Table 4. Adaptation of acquisition dose to patient's weight.

Body weight (kg)	kV	mAs_{ref}
<50	80	1.5 mAs/kg
50–70	100	120
70–90	100 (120)	150 (100)
>90 or arms aside	120 (140)	150

can be further optimised adjusting kilovoltage and milliamperage according to the patient's weight (Table 4).

Mostly, 120 kVp and about 100 mAs are chosen as standard exposure parameters for pulmonary CTA with higher kilovoltage settings (i.e. 140 kVp) for obese patients. With the introduction of multidetector row CT, lowering the kVp from the standard 120–100 kVp reduces the $CTDI_{vol}$ by a factor of 2–3. Due to the absorption characteristics of iodium, the intravascular opacification is increased by some 70–100 HU with lower kVp compensating for the generally higher image noise. The latter is the reason that lowering the kVp is only recommended for children and low and normal weighted patients, though not for adipose patients [21-23]. The resulting effective dose will be in the range of 1.5–2 mSv. A tube voltage of 80 kVp holds an even higher potential for dose savings in slim individuals and children, with an effective dose under 1.5 mSv. In a recent study, simulated low-dose CT scans, generated by the superimposition of computer-calculated noise comparable with low mAs setting, were compared with a reference standard of 90 mAs in non-obese patients [24]. No statistically significantly differences in the visualisation of peripheral PE and inter- and intraobserver agreement were found.

The resulting effective dose depends very much on the chosen scan length. Following the equation dose–length product (DLP) = scan length × CTDIvol in mGy × cm, the effective dose *E* for a total chest CTA (30 cm), is about 40% of the $CTDI_{vol}$ (or, the conversion factor is 0.015 mSv per mGy × cm for a total chest CTA). If the scan range can be reduced, the effective dose decreases proportionally. As a consequence, the effective dose may vary substantially despite the fact that the same mAs and kVp settings had been chosen. The European Guidelines for Quality in Computed Tomography EUR 16262 suggests a maximum effective dose level for the chest of 9 mSv. In our experience, a $CTDI_{vol}$ of 5–7 mGy for a standard size patient (70 kg, 170 cm) is sufficient. A recent comparative study between multidetector CTA and pulmonary DSA found a mean effective dose of 4.2 mSv (range 2.2–6.0 mSv) for CTA and 7.1 mSv (range 3.3–17.3 mSv) for DSA [25].

CT DURING PREGNANCY

The prevalence of acute PE is not only five times higher during pregnancy; it is also the most important cause of maternal death. Whether CTA or Q-scintigraphy should be the first image modality of choice with young women is still a matter of debate and should also be chosen on the basis of availability and experience to ensure an optimal diagnostic outcome of the examination. The most important disadvantage of CT is the high breast dose, which is even more harmful to the more sensitive breast tissue during pregnancy (up to 40 times higher as compared to Q-scintigraphy). This is an important argument to use perfusion scintigraphy in patients with normal chest radiographs and a low risk of overlying lung disease. Arguments in favour of CT are the more secure diagnostic yield of both PE and alternative diagnosis. International recommendations also include the ultrasound of the leg veins as a primary diagnostic approach, because secure positive findings justify treatment without further irradiation of the chest.

Although not routinely used, shielding is considered a good option to reduce the dose to the breasts. Indeed, Geleijns et al. showed that a dose reduction of 30% to the breast and 15% to the lungs could be obtained using bismuth shielding [26]. However, bismuth shields cause beam hardening artefacts and increase image noise in the chest. It therefore seems that the reduction in radiation dose obtained with bismuth shielding could be achieved more efficiently by a reduction of tube current (see below).

It is noteworthy that – based on phantom studies – during the complete duration of pregnancy, the foetal dose with CT is lower or of the same magnitude as with Q-scintigraphy as the radiopharmaceutical is excreted by the kidneys and accumulates in the bladder [27-29]. However, it should also be noted that – whatever technique is used – the foetal radiation dose is less than 1 mGy and therefore remains far below the upper limit of 50 mGy for induct.

Recent studies have shown that the risk of yielding a suboptimal CTA image quality increases significantly during pregnancy when standard CTA acquisition protocols are used [30-32]. On one hand, this is the result of reduction of vascular contrast enhancement due to the hyperdynamic circulation during pregnancy that occurs already in the early phase of pregnancy; on the other hand, there is a change in breathing pattern with pregnant women with an increased risk for Valsalva manoeuvre. To improve the image quality and keep the radiation as low as possible, an adaptation of the protocol is warranted [33,34]. Parameters that can be adapted to optimise the protocol for pregnant patients are given in Table 5.

Table 5. Methods to improve the image quality and to reduce the radiation dose of CTA for acute PE during pregnancy.

Improvement of image quality
Short scan duration (choosing the fastest scanner)
High iodine influx (↑ increase of flow and / or iodine concentration, e .g. 6 ml /s and 400 iodium/ml)
No maximal inspiration or even shallow breathing
Reduction of radiation doses
Reduction of tube current
Reduction of tube voltage (e.g., 100 kVp)
Reduction of z-axis (limited scan range)
Increase in pitch (1.5–2)
Increase in collimation (1.5 mm)
Standard use of dose modulation

CTA AS ONE-STOP SHOP

Multiple publications and a lot of debate have recently been dealing with CTPA as part of a one-stop diagnostic approach. Two applications are under discussion: one includes the venous phase scan of the abdomen and lower extremities [3,35,36] and the second refers to an appropriate CT protocol that allows assessment of both, cardiac and non-cardiac vascular causes of acute chest pain in emergency department patients [14].

Thromboembolic disease is considered a systematic disease with thrombi sitting in the lower leg veins and/or emboli sitting in the pulmonary arteries, both requiring equivalent therapeutic consequences. Therefore, it was suggested to examine also the deep venous system starting at the subdiaphragmatic level or the pelvis and examining the lower extremities down to the popliteal fossa within the same CT examination without a second injection of contrast medium [33,34]. However, to optimise the opacification of the veins, more contrast (120–150 ml) has to be injected. A discontinuous axial scanning technique with acquisition of slices every 2–5 cm, or a continuous spiral scanning technique using relatively thick slice collimation have been applied and yielded comparable results [36]. The additional radiation dose is the main issue of concern; it amounts to 2.3–8.3 mSv depending on the protocol applied and has to be weighed against the potential diagnostic benefit [35,36]. Substantial numbers of patients – in the PIOPED II trial 3–5% – have to be exposed to radiation to find those patients who have DVT despite a negative pulmonary CTA and therefore require treatment [3]. This means that the vast majority of patients will be exposed to additional radiation without

clinical consequences. For that reason, at least in Europe, the deep venous system is in most hospitals not routinely included in the CTA examination. If applied (e.g. on ICU patients that are difficult to examine with ultrasound, in patients with suspicion for thrombi in the iliac veins and the inferior vena cava, that are more difficult to sufficiently examine with ultrasound, or on patients with a high clinical suspicion and a CTA negative for PE, however this requires on-site assessment of the pulmonary CTA), low-dose scanning should be mandatory and may be achieved by thick sections, low kV and low mAs settings. It has to be noted that the interobserver agreement for evaluation of the venous system in CTA examinations has been found to be lower than for the pulmonary arterial scan [37]. Inhomogeneous flow, suboptimal opacification due to inflow problems (i.e. severe atherosclerotic disease or decreased cardiac function) or artefacts caused by orthopedic hardware or vascular calcification can create pseudo-filling defects and may lead to false negative results or indeterminate scan results [37,38].

Another so-called one-stop shop approach refers to the triple rule-out protocol in patients with acute chest pain that potentially may have an aortic dissection, PE or a coronary disease as underlying reason for their symptoms. A recent prospective study using the dual-source CT scanning technique with retrospective ECG-gating demonstrated excellent image quality and a high diagnostic yield for cardiovascular disease [39]. If CTA of the pulmonary arteries is performed within a triple rule-out protocol, retrospective ECG-gating is needed. Scan parameters (140 kVp, 350–500 mAs) and injection protocols (120–150 ml of contrast medium, region-of-interest in the ascending aorta) are accordingly adapted. An efficient implementation of such a protocol requires special hardware and software conditions and expertise that may be difficult to provide in an on-call setting. The majority of used CT scanner types and the extension of the scan volume coverage result in a considerably higher radiation dose. The discussion whether a triple rule-out protocol is safe, cost-effective, improves clinical decision making and which patient category may benefit from this protocol is still ongoing and beyond the scope of this article.

Image post-processing and evaluation

Window settings

Image interpretation is typically performed using both a soft tissue or mediastinal window setting (window width (WW) = 400 HU; window level (WL) = 30–40 HU) and a pulmonary parenchyma window setting (WW = 1500 HU; WL = –800 to –600 HU). Some authors describe a PE specific window (WW = 700 HU;

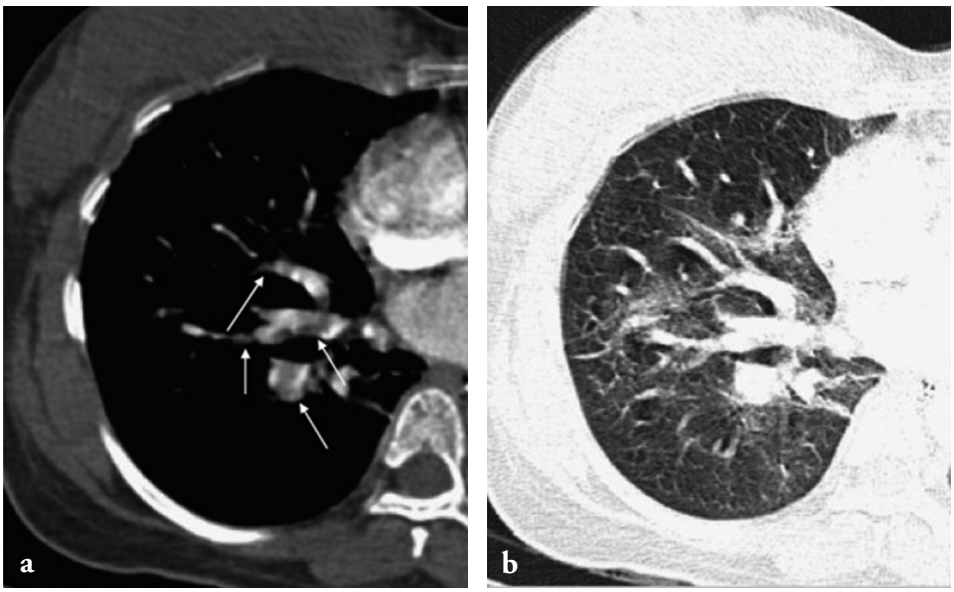

Figure 3. CT scan in mediastinal window setting (**a**) shows a filling defect in a segmental lower lobe pulmonary artery (arrow) that was found to be a breathing artefact on pulmonary window setting (**b**).

WL = 100 HU). With use of fixed standard window settings, small PE may be obscured by dense contrast medium. Therefore, an individual adaptation of the window setting is advantageous, corresponding to a WW of slightly less than twice the mean attenuation in the main pulmonary artery and a WL of about half the mean attenuation in the pulmonary artery [40]. A poor enhancement is often more difficult to overcome, narrowing the window width and window level setting may result in a more confident interpretation. Parenchymal window settings are indispensable for the assessment of motion artefacts as underlying reason for intravascular contrast inhomogeneities (Figure 3) and for differentiation of pulmonary arteries from mucous filled bronchi or from venous structures which, in the early phase of scanning, may be unenhanced.

In addition to the assessment of the axial slices, multiplanar reformats (MPR) or curved planar reformats (CPR), e.g. along the long axis of the vessel of interest and perpendicular to its lumen are used as problem solving tools: they help to properly assess pulmonary arteries that are oriented obliquely or perpendicular to the imaging plane, to distinguish between central clots and perivascular lymphatic tissue (Figure 4) or to differentiate between pulsation artefacts and real emboli. Findings made in MPR and CPR should be correlated with findings in the axial plane.

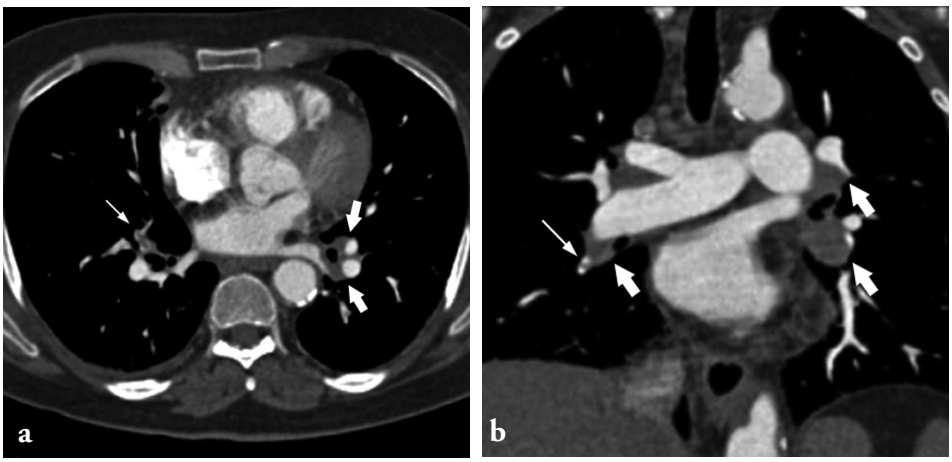

Figure 4. Axial (**a**) and coronal (**b**) MPR showing hilar lymphadenopathy (large arrows) and a small pulmonary embolus (small arrow).

Maximum intensity projections (MIP)

Maximum intensity projections (MIP) are an excellent tool for providing angiography-like images. Sliding thin slab MIPS (3–10 mm thickness) were found to improve delineation of small peripheral vessels (Figure 5). Small peripheral clots will appear as vessel segments that are less enhanced (i.e. have soft-tissue density) than neighbouring vessels of the same size. Choosing sufficiently thin slabs and adjustment of the window settings decrease the risk to miss intravascular hypodense clots that may otherwise be hidden by a surrounding high density contrast medium (Figure 6).

Computer-assisted detection

Interpreting a CTPA for the detection of PE is a demanding visual task due to the high number of images to be evaluated, the various PE look-alikes and human factors such as attention span and experience. Computer-assisted detection (CAD) software has therefore been developed to help readers to improve detection performance, increase reader homogeneity and potentially decrease reading time.

Multiple manufacturers developed a CAD algorithm the majority of which is still under evaluation and not yet FDA approved or introduced into clinical routine. Most studies so far focused on the standalone performance and found a wide range of sensitivities on a per lesion basis between 31% and 100%. The average number of FP findings per scan varied in these studies between 0.93 and 14.4 per scan. This variation is not only caused by differences of the CAD algorithms, but largely due to the different sizes and composition of study

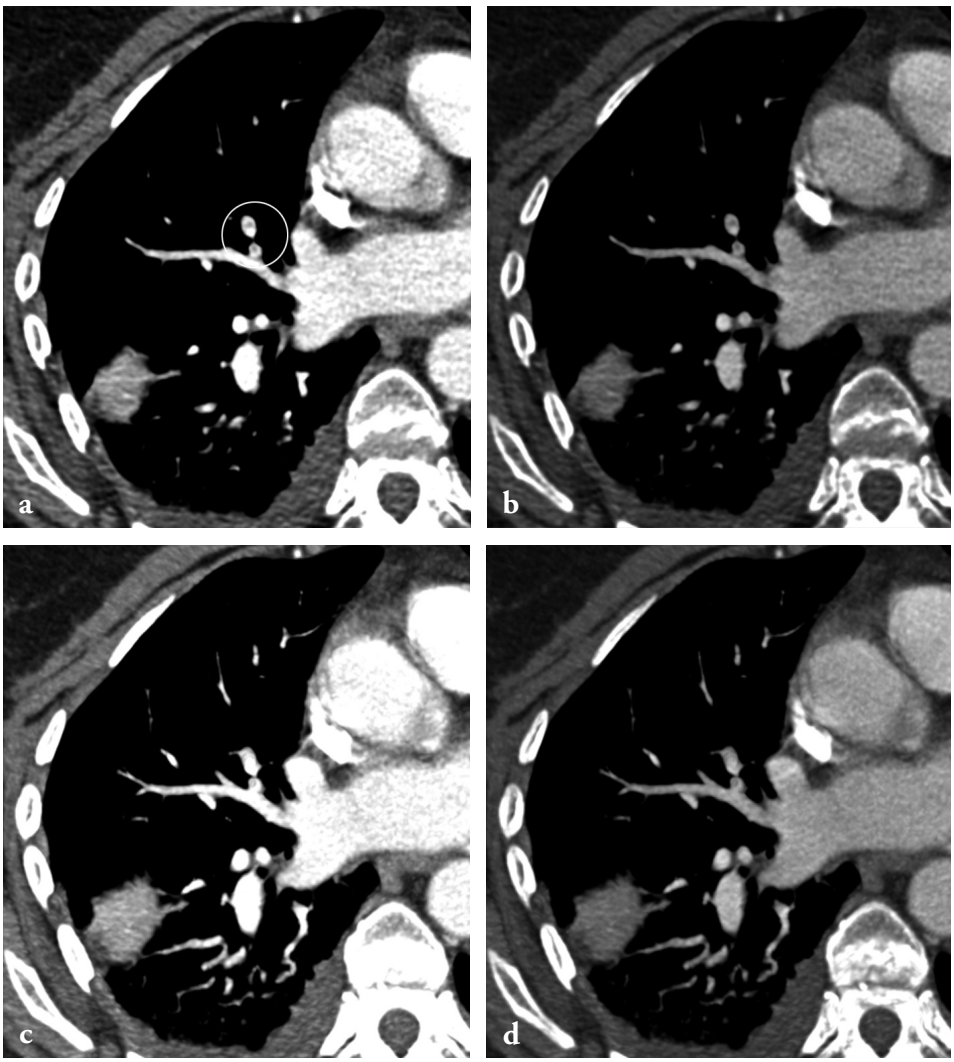

Figure 5. The effect of slice thickness and different window settings on the detection of a small pulmonary embolus (circle): note the superiority of thinner slices for both MPR and MIP. (**a**) 1 mm slice thickness MPR, window width/window level (WW/WL) = 400/40, (**b**) 1 mm slice thickness MIP, WW/WL = 600/120, (**c**) 3 mm slice thickness MPR, WW/WL = 400/40, (**d**) 3 mm slice thickness MIP, WW/WL = 600/120.

groups. Only two studies published so far used more than 100 non-selected cases [41,42], while other studies used relatively small numbers of cases or selected cases with exclusively good image quality. Though in routine, diagnosis of PE is made on a per patient basis, fewer results are known on patient basis: here,

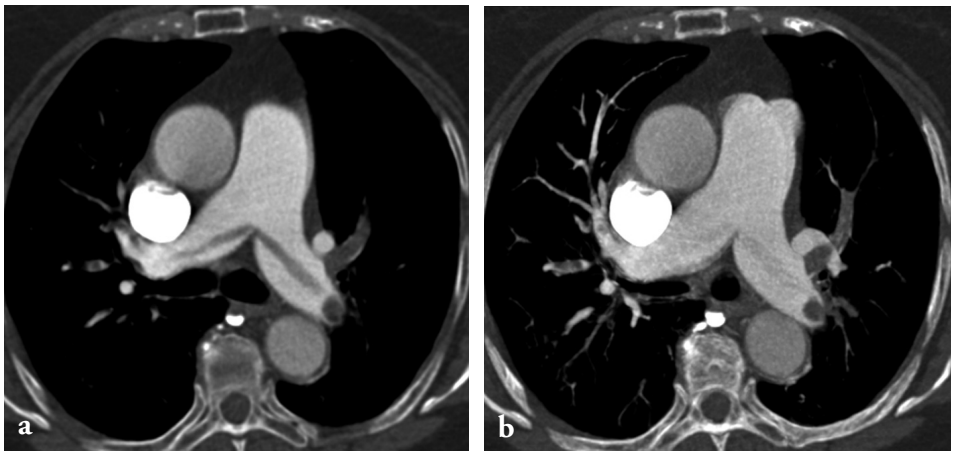

Figure 6. CT scan in mediastinal window setting reveals a central pulmonary embolus on a 5 mm axial MPR reconstruction (**a**) and a 5 mm axial MIP (**b**). It is clear that the central part of the thrombus is less easily visualised on the thick MIP reconstruction.

the sensitivities varied between 54%, 86% and 94%, respectively [41-43]. Whether CAD software eventually stand up to its goals and improve and homogenise radiologists' performance in clinical practice will depend not only on the accuracy of the system itself, but also on a number of reader related factors such as experience, confidence into the system and the ability to differentiate true from false positive candidate lesions.

In summary, studies so far show some promising results for the use of CAD algorithms. However, CAD systems still need to be substantially improved with respect to their specificity before being ready for use in clinical routine and prospective studies in non-selected patient groups are necessary to study the impact of CAD on readers' diagnosis and thus eventually on patient care.

Thrombus load and right ventricular function

With regard to prognosis and therapy, patients with PE are traditionally classified in two groups: those who are hemodynamic insufficient (i.e. hypotension, cardiac arrest or shock) and those who are stable. The latter group has a good prognosis with low risk of PE-related morbidity and mortality when anticoagulant therapy is promptly installed, in contrast to the former group who has a high risk of PE-related mortality and in whom thrombolytic therapy is indicated. In addition, various diagnostic and laboratory tests have been investigated to

define to which extent subgroups of patients that, although hemodynamic stable, may also benefit from more aggressive fibrinolytic therapy or surgery (embolectomy) or vice versa which patient groups may equally profit from a less aggressive therapy.

Several studies have focused on right ventricular (RV) function, dilatation of the pulmonary trunk, bowing of the intraventricular septum and/or the extent of vascular obstruction as an independent prognostic factor in the outcome of patients with PE who are hemodynamically stable [44]. The presence of right ventricular dysfunction assessed by echocardiography was found to be a predictor of less favourable outcomes even in normotensive patients. Although cardiac function is traditionally assessed by ultrasound, the presence of RV dilatation (RV:LV diameter >1) assessed on axial CT (Figure 7) seems also to be directly related to mortality in PE patients. Since CTA has become the first method of choice as an imaging tool in the work-up of PE, information on RV function is automatically available and may supersede the need of an additional echocardiography. In addition, RV dilatation assessed on axial CT images was found to be comparable to reconstructed 4-chamber views, obviating the need of additional reconstructions [45].

In addition to these cardiac morphologic findings, the thrombus load and the degree of arterial obstruction can also be assessed with CT. Various scoring systems, based on either the conventional angiography scoring system as defined by Miller (Bankier) or CT-derived (Mastora, Qanadli), have been proposed to

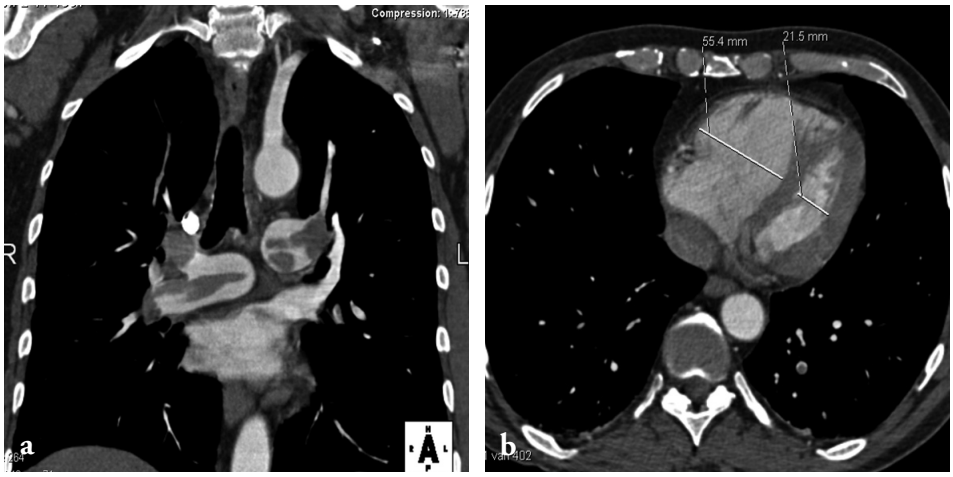

Figure 7. A 59-year-old man with a suspicion of pulmonary embolism. Coronal MPR demonstrating bilateral massive pulmonary embolism (**a**) with dilatation of the right ventricle and bowing of the interventricular septum (**b**).

assess the thrombus load and degree of arterial obstruction in patients with acute PE. The relation between clot burden on one side and either short-term survival or RV dysfunction on the other side is still controversial. Furthermore, these scores so far never found their entrance into clinical routine because they are very time consuming. With an effective "automatic detection of thrombi", this limitation may be overcome in the future. Contradicting results have also been published on the role of other findings such as dilatation of the pulmonary trunk and the shape of the intraventricular septum as independent predictors of short-term mortality.

With the advent of the ultra-fast scanners, it is now possible to determine not only morphologic cardiac findings but also functional parameters such as RV ejection fraction [46]. It has to be noted that in the future, the diagnostic value of simultaneous assessment of right ventricular function will be determined by the fact whether therapeutic consequences will be based upon it. Although the relation between RV dysfunction and worse outcome has been confirmed in several studies and the arterial obstruction index may be an independent predictor (although its role is still controversial), it still remains to be determined whether the subgroup of patients with RV dysfunction or significant arterial obstruction truly may benefit from more intensive treatment with thrombolytics.

Summary

Due to the development of faster MDCT techniques and optimising the contrast medium injection and scan protocols, the quality of the CTA for the detection of PE has considerably increased and consequently reduced the percentage of non-diagnostic results. Inadequate vascular enhancement is one of the most important causes for a non-diagnostic scan result. Techniques to improve the vascular enhancement include lowering the kV or increasing the amount of iodine injected per second. With the development of the newer generation scanners, the radiation dose could be substantially decreased. Further dose reduction is possible by reducing the scan range, reducing the kVp and using weight adapted scan parameters. Special adaptations of the scan and contrast injection protocols are recommended with young patients and pregnant women.

In addition, CTA can be used as a diagnostic one-stop-shop approach either in an emergency setting to assess both cardiac and non-cardiac vascular causes of acute chest pain, or to evaluate the veins of abdomen and lower extremities to detect deep vein thrombosis in addition to CTA of the chest.

With a dual-source scanner, lung perfusion can be visualised directly which may give additional functional information when PE is present. Computer-

assisted diagnosis software packages have been developed in order to increase the detection of PE and decrease the reading time. These new techniques are very promising but their value within the diagnostic algorithm for acute PE still needs to be determined

Reference List

1. Remy Jardin M, Remy J, Wattinne L et al. Central pulmonary thromboembolism: diagnosis with spiral volumetric CT with the single-breath-hold technique—comparison with pulmonary angiography.Radiology 1992; 185:381–387.
2. Jones SE and Wittram C. The indeterminate CT pulmonary angiogram: imaging characteristics and patient clinical outcome. Radiology 2005; 237:329–337.
3. Stein PD, Fowler SE and Goodman LR et al. Multidetector computed tomography for acute pulmonary embolism. N Engl J Med 2006; 354:2317–2327.
4. Ghaye B, Szapiro D, I. Mastora et al. Peripheral pulmonary arteries: how far in the lung does multi-detector row spiral CT allow analysis? Radiology 2001; 219:629–636.
5. Patel S, Kazerooni EA, Cascade PN. Pulmonary embolism: optimization of small pulmonary artery visualization at multi-detector row CT. Radiology 2003; 227:455–460.
6. Schoepf UJ, Holzknecht N, Helmberger TK et al. Subsegmental pulmonary emboli: improved detection with thin-collimation multi—detector row spiral CT. Radiology 2002; 222:483–490.
7. Bankier AA, O'Donnell CR, Boiselle PM. Quality initiatives: respiratory instructions for CT examinations of the lungs: a hands-on guide. Radiographics 2008; 28:919–931.
8. Gosselin MV, Rassner UA, Thieszen SL et al. Contrast dynamics during ct pulmonary angiogram. J Thorac Imaging 2004; 19:1–7.
9. Pontana F, Faivre JB, Remy-Jardin M et al. Lung perfusion with dual-energy multidetector-row CT (MDCT): feasibility for the evaluation of acute pulmonary embolism in 117 consecutive patients. Acad Radiol 2008; 15:1494–1504.
10. Thieme SF, Johnson TRC, Lee C et al. Dual-energy CT for the assessment of contrast material distribution in the pulmonary parenchyma. Am J Roentgenol 2009; 193:144–149.
11. Fink C, Johnson TR, Michaely HJ et al. Dual-energy CT angiography of the lung in patients with suspected pulmonary embolism: initial results (Dual-Energy-CT-Angiografie der Lunge bei Patienten mit Verdacht auf Lungenembolie: Erste Ergebnisse). Rofo (Fortschr Röntgenstr) 2008; 180:879–883.
12. Zhang LJ, Chai X, Wu SY et al. Detection of pulmonary embolism by dual energy CT: correlation with perfusion scintigraphy and histopathological findings in rabbits. Eur Radiol 2009; 19:2844–2854.
13. Marten K, Engelke C, Obenauer S et al. Diagnostic performance of retrospectively ecg-gated multislice CT of acute pulmonary embolism, Diagnostischer Stellenwert der retrospektiven EKG-Triggerung in der Mehrschicht-Spiral-CT der akuten Lungenembolie. Fortschr R"ntgenstr 2003; 175:1490–1495.
14. White CS, Kuo D, Kelemen M et al. Chest pain evaluation in the emergency department: can MDCT provide a comprehensive evaluation? Am J Roentgenol 2005; 185:533–540.

15. Schoellnast H, Deutschmann HA, Fritz GA et al. MDCT angiography of the pulmonary arteries: influence of iodine flow concentration on vessel attenuation and visualization. Am J Roentgenol 2005; 184:1935–1939.
16. Haage P, Schmitz-Rode T, Hubner D et al. Reduction of contrast material dose and artefacts by a saline flush using a double power injector in helical CT of the thorax. Am J Roentgenol 2000; 174:1049–1053.
17. Coche EE, Hammer FD, Goffette PP. Demonstration of pulmonary embolism with dynamic gadolinium-enhanced spiral CT. Eur Radiol 2001; 11:2306–2309.
18. Remy-Jardin M, Bahepar J, Lafitte JJ et al. Multi-detector row CT angiography of pulmonary circulation with gadolinium-based contrast agents: prospective evaluation in 60 patients1. Radiology 2006; 238:1022–1035.
19. Greess H, Wolf H, Baum U et al. Dose reduction in computed tomography by attenuation-based online modulation of tube current: evaluation of six anatomical regions. Eur Radiol 2000; 10:391–394.
20. Mastora I, Remy-Jardin M, Suess C et al. Dose reduction in spiral ct angiography of thoracic outlet syndrome by anatomically adapted tube current modulation. Eur Radiol 2001; 11:590–596.
21. Heyer CM, Mohr PS, Lemburg SP et al. Image quality and radiation exposure at pulmonary CT angiography with 100- or 120-kvp protocol: prospective randomized study 1. Radiology 2007; 245: 577–583.
22. Schueller-Weidekamm C, Schaefer-Prokop CM, Weber M et al. CT angiography of pulmonary arteries to detect pulmonary embolism: improvement of vascular enhancement with low 23. Sigal-Cinqualbre AB, Hennequin R, Abada HT et al. Low-kilovoltage multi-detector row chest CT in adults: feasibility and effect on image quality and iodine dose. Radiology 2004; 231:169–174.
24. Tack D, De Maertelaer V, Petit W et al. Multi-detector row CT pulmonary angiography: comparison of standard-dose and simulated low-dose techniques. Radiology 2005; 236:318–325.
25. Kuiper J, Geleijns J, Matheijssen N et al. Radiation exposure of multi-row detector spiral computed tomography of the pulmonary arteries: comparison with digital subtraction pulmonary angiography. Eur Radiol 2003; 13:1496–1500.
26. Geleijns J, Salvadó Artells M, Veldkamp W et al. Quantitative assessment of selective in-plane shielding of tissues in computed tomography through evaluation of absorbed dose and image quality. Eur Radiol 2006;16.
27. Winer-Muram HT, Boone JM, Brown HL et al. Pulmonary embolism in pregnant patients: fetal radiation dose with helical CT. Radiology 2002; 224:487–492.
28. Patel SJ, Reede DL, Katz DS et al. Imaging the pregnant patient for nonobstetric conditions: algorithms and radiation dose considerations. Radiographics 2007; 27:1705–1722.
29. Doshi SK, Negus IS, Oduko JM. Fetal radiation dose from ct pulmonary angiography in late pregnancy: a phantom study. Br J Radiol 2008; 81:653–658.
30. Andreou A, Curtin J, Wilde S et al. Does pregnancy affect vascular enhancement in patients undergoing ct pulmonary angiography? Eur Radiol 2008; 18:2716–2722.

31. King-Im JU, Freeman S, Boylan T et al. Quality of CT pulmonary angiography for suspected pulmonary embolus in pregnancy. Eur Radiol 2008; 18:2709–2715.
32. Ridge CA, McDermott S, Freyne BJ et al. Pulmonary embolism in pregnancy: comparison of pulmonary CT angiography and lung scintigraphy. Am J Roentgenol 2009; 193:1223–1227.
33. Pahade JK, Litmanovich D, Pedrosa I et al. Imaging pregnant patients with suspected pulmonary embolism: what the radiologist needs to know. Radiographics 2009; 29:639–654.
34. Schaefer-Prokop C, Prokop M. CTPA for the diagnosis of acute pulmonary embolism during pregnancy, Eur Radiol 2008; 18:2705–2708.
35. Cham MD, Yankelevitz DF, Henschke CI. Thromboembolic disease detection at indirect ct venography versus ct pulmonary angiography. Radiology 2005; 234:591–594.
36. Ghaye B, Dondelinger RF. Non-traumatic thoracic emergencies: CT venography in an integrated diagnostic strategy of acute pulmonary embolism and venous thrombosis. Eur Radiol 2002; 12:1906–1921.
37. Garg K, Kemp JL, Russ PD et al. Thromboembolic disease: variability of interobserver agreement in the interpretation of CT venography with CT pulmonary angiography. AJR Am J Roentgenol 2001; 176:1043–1047.
38. Taffoni MJ, Ravenel JG, Ackerman SJ. Prospective comparison of indirect ct venography versus venous sonography in ICU patients. Am J Roentgenol 2005; 185:457–462.
39. Schertler T, Frauenfelder T, Stolzmann P et al. Triple rule-out CT in patients with suspicion of acute pulmonary embolism: findings and accuracy1. Acad Radiol 2009; 16:708–717.
40. Bae KT, Mody GN, Balfe DM et al. CT depiction of pulmonary emboli: display window settings. Radiology 2005; 236:677–684.
41. Walsham AC, Roberts HC, Kashani Hm et al. The use of computer-aided detection for the assessment of pulmonary arterial filling defects at computed tomographic angiography. J Comput Assist Tomogr 2008; 32:(6):913–918.
42. 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 2009; 20(4):801-6.
43. Maizlin ZV, Vos PM, Godoy MC et al. Computer-aided detection of pulmonary embolism on CT angiography: initial experience. J Thorac Imaging 2007; 22:324–329.
44. Ghaye B, Ghuysen A, Bruyere PJ et al. Can CT pulmonary angiography allow assessment of severity and prognosis in patients presenting with pulmonary embolism? What the radiologist needs to know. Radiographics 2006; 26:23–40.
45. Kamel EM, Schmidt SM, Doenz FM et al. Computed tomographic angiography in acute pulmonary embolism: do we need multiplanar reconstructions to evaluate the right ventricular dysfunction? J Comput Assisted Tomogr 2008; 32:438–443.
46. Doğan H, Kroft LJ, Huisman MV et al. Right ventricular function in patients with acute pulmonary embolism: analysis with electrocardiography-synchronized multidetector row CT. Radiology 2007; 242:78–84.

Part 2

Standalone performance of a computer-assisted detection prototype for the detection of acute pulmonary embolism

Chapter 3

Wittenberg R
Peters JF
Sonnemans JJ
Prokop M
Schaefer-Prokop CM

Eur Radiol 2010 Apr;20(4):801-6.

Computer-assisted detection of pulmonary embolism: evaluation of pulmonary CT angiograms performed in an on call setting

Abstract

Purpose

The purpose of the study was to assess the standalone performance of computer-assisted detection (CAD) for evaluation of pulmonary CT angiograms (CTPA) performed in an on call setting.

Methods

In this institutional review board-approved study, we retrospectively included 292 consecutive CTPA performed during night shifts and weekends over a period of 16 months. Original reports were compared with a dedicated CAD system for pulmonary embolism (PE). A reference standard for the presence of PE was established using independent evaluation by two readers and consultation of a third experienced radiologist in discordant cases.

Results

Original reports had described 225 negative studies and 67 positive studies for PE. CAD found PE in seven patients originally reported as negative but identified by independent evaluation: emboli were located in segmental (n=2) and subsegmental arteries (n=5). The negative predictive value (NPV) of the CAD algorithm was 92% (44/48). On average there were 4.7 false positives (FP) per examination (median 2, range 0–42). In 72% of studies ≤5 FP were found, 13% of studies had ≥10 FP.

Conclusion

CAD identified small emboli originally missed under clinical conditions and found 93% of the isolated subsegmental emboli. On average there were 4.7 FP per examination.

Introduction

Acute pulmonary embolism (PE) is one of the most frequent potentially fatal diseases. Unfortunately, because of its unspecific clinical symptoms, PE is difficult to diagnose clinically. In two-thirds of the cases the diagnosis is missed if no additional diagnostic procedures are carried out [1-3].

The diagnostic imaging method of choice for the detection of PE in most institutions is contrast-enhanced pulmonary computed tomography angiography (CTPA) [4-8]. However, each CTPA examination produces on average 300–500 axial images. Thin sections need to be reviewed because it is known that small emboli can be missed on thick sections [9]. Meticulous review of all CT slices is therefore time-consuming and requires a high level of attentiveness. The prevalence of PE in patients sent to CTPA is moderate to low, in the range of 10–35% according to current literature [10,11]. The chance of missing small emboli increases with time pressure and anatomical and technical complexity and decreases as readers become more experienced [12,13].

Computer-assisted detection (CAD) algorithms have been developed to help exclude PE and to improve the detection performance of observers. The purpose of this study was to assess the standalone performance of a CAD prototype on evaluation of pulmonary CT angiograms (CTPA) performed in an on call setting. Performance of the CAD algorithm as a standalone system was compared with a reference standard and the original reports.

Materials and methods

Patient selection

In this institutional review board-approved study, we retrospectively included all 292 consecutive CTPA studies performed in a university hospital during night shifts and weekends over a period of 16 months between January 2007 and April 2008. All patients had been referred to the radiology department for CTPA because of suspected acute PE. Fourteen patients were excluded from further evaluation for the following two reasons: six CTPA examinations because of streak artefacts based on non-elevated arm positions and metallic material that made a diagnostic evaluation of the imaging impossible and eight examinations because the CAD algorithm did not work for these data sets. In two of these cases, the CAD algorithm failed because the patient had a pneumothorax on the left side or the left lung had been surgically removed. The other six patients had a trachea tube for respiratory ventilation resulting in a connection between extrathoracic air and intrapulmonary air and subsequent

failure of the segmentation of the trachea and lungs.

The final study group therefore consisted of 278 patients: 138 male, 140 female, mean age 57 years (range 18–88). There were 133 inpatients (13 patients came from the intensive care unit), 5 patients from the outpatient clinic and 140 patients submitted to the emergency unit. In total, 23% studies were originally reported as positive for PE and 77% as negative.

CT technique

All CTPA examinations were acquired using 16- or 64-detector-row CT; 178 patients were examined by 16-detector-row CT (MX 8000 IXDT or Brilliance-16, Philips Medical Systems, Cleveland, OH) and 100 patients underwent 64-detector-row CT (Brilliance-64, Philips Medical Systems, Cleveland, OH). All CT images were acquired in a caudo-cranial direction from the level of the diaphragm to the lung apices within a single breath hold. A standard PE protocol was applied. The Stellant Dual CT Injector (Medrad Europe BV, Beek, The Netherlands) was used for intravenous bolus injection.

The 64-slice CTPA were obtained with 120 kV, 100 mAs, 64×0.625 mm collimation, rotation time 0.4 s and pitch 1.172. All images were reconstructed with 0.9-mm slice thickness and 0.45-mm overlap. Patients received 90 ml of i.v. contrast medium (Ultravist 300, Schering, Berlin, Germany) injected at a flow rate of 5.0 ml/s and followed by a 30-ml NaCl chaser bolus. The 16-slice CTPA were obtained with 90 kV, 180 mAs, 16×0.75 mm collimation, rotation time 0.5 s, and pitch 1.188. All images were reconstructed with 1.0-mm slice thickness and 0.5-mm overlap. Patients received 90 ml of i.v. contrast medium (Ultravist 300), injected at a flow rate of 4.0 ml/s and followed by a 40-ml NaCl chaser bolus. For both techniques, a bolus tracking method was applied with the ROI in the pulmonary trunk. The threshold to start data acquisition was set at 150 HU and a start delay of 8 s after reaching the trigger threshold was used.

Original reports had been based on the evaluation of thin axial images. Multiplanar reconstructions and maximum intensity projections (MIP) were used at the discretion of the interpreting resident or radiologist.

Data collection and analysis

All examinations were analysed by CAD prototype software (Philips Healthcare, Best, The Netherlands). For the analysis of the CAD lesions, a reference standard was established by independent evaluation by two readers. A researcher specially trained in reading PE studies (RW) and an experienced chest radiologist (>15 year experience, CMS) evaluated all datasets separately. In the case of discordant findings between these two readers, a third experienced chest radiologist (>15 years

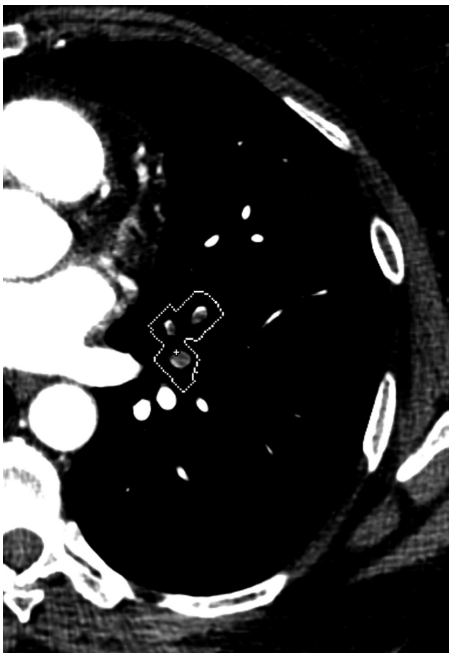

Figure 1. Classis embolus found by CAD

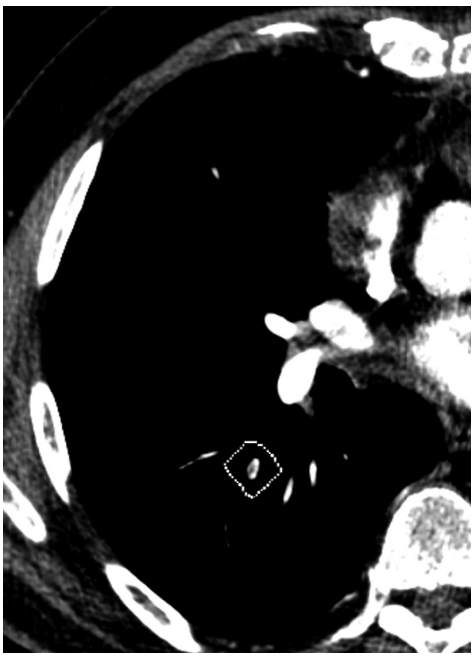

Figure 2. Subsegmental embolus missed by primary readers

experience, MP) was consulted. The readers were instructed to digitally mark each intravascular thrombus with regard to its anatomical location (e.g. pulmonary lobe) and the anatomical level of its proximal end (central, lobar, segmental and subsegmental). Standard nomenclature, derived from Boyden [14] and from Jackson and Huber [15], was used to identify the segmental and subsegmental structures. Care was taken that thrombi with continuous extension in various branching vessels were counted as a single lesion and that small subsegmental emboli were included.

In a second step, the CAD findings were compared with our reference standard and considered as true positive, false negative or false positive (Figure 1,2,3). To define a CAD finding as true positive that had not primarily been seen by one of the readers, the presence of an intravascular contrast defect had to be securely confirmed by the other readers.

Principles of the CAD algorithm

CAD is a fully automated PE detection prototype system (Philips Healthcare, Best, The Netherlands). It starts from automatic segmentation of the trachea and lungs. The resulting lung volume after morphological processing contains

Figure 3. False positive candidate lesion found by CAD due to motion artefacts

both tissue and air. Subsequently, pulmonary vessels are segmented within this volume, starting with segmentation of all structures above -100 HU. Second, a vessel-tracking algorithm is applied to extract the pulmonary vessels [16]. After pruning the vessels, vessel cross-sectional images perpendicular to the centrelines are computed. Grey value analysis of these cross-sectional images is applied to find candidate locations for PE. No PE are found outside the lung segmentation. Features for identifying PE include stretches of vessels that are completely occluded as well as areas that contain contrast-to-tissue transitions. Candidate lesions that the CAD identifies as potential PE are clustered and presented to the reader, indicated by an ROI around the affected vessel. The processing requires about 30 s per examination and is carried out automatically in the background during initial evaluation by the radiologists. Candidate lesions are then shown only on demand.

Statistical analysis

Statistical analysis was performed with SPSS (version 15.0 statistics UK). Sensitivity for the detection of PE by CAD was assessed on a per-patient as well as on a per-lesion basis, the latter with respect to the main, lobar, segmental

Table 1. Diagnosis according to the original reports versus the reference standard

	Standard positive	Standard negative	Total
Original report positive	61	4	65
Original report negative	7	206	213
Total	68	210	278

Table 2. Diagnosis based on the CAD findings compared with the reference standard

	Standard positive	Standard negative	Total
CAD positive	64	166	230
CAD negative	4	44	48
Total	68	210	278

and subsegmental levels. Specificity and the negative predictive value (NPV) of an examination without any candidate lesions found by the CAD algorithm were calculated.

Results

Distribution of emboli and sensitivity per embolus

Our reference standard determined that 68 out of 278 patients were positive for PE. A total of 377 emboli were detected in the 68 patients. Twenty-three emboli were localised in the main pulmonary arteries, 41 in lobar, 80 in segmental and 233 in the subsegmental arteries. In 37 patients, there were emboli extending into the main and/or lobar arteries. Thirty-one patients had only segmental and/or subsegmental emboli. Fourteen patients had isolated subsegmental emboli. The per-embolus sensitivity of the CAD as a standalone system was 87, 78, 79 and 61% respectively for emboli in the main, lobar, segmental and subsegmental arteries.

Patientbased evaluation

CAD correctly identified PE in seven patients originally reported as negative. In five patients, CAD found a solitary embolus in a segmental (n=2) or subsegmental (n=3) artery. In two patients, CAD found two subsegmental emboli. On a per-patient basis, the sensitivity of the CAD system was 94% (64/68) and the

specificity 21% (44/210). CAD correctly identified 93% (13/14) of the patients with isolated subsegmental emboli.

In four patients, we found no emboli although the original reports had been positive. In 21% of the negative cases, there were no FP findings by CAD. The negative predictive value (NPV) of an examination without any candidate lesion found by the CAD algorithm was 92% (Tables 1 and 2).

For the remaining 278 examinations, the CAD algorithm showed on average 4.7 FP lesions per examination (median 2, range 0–42) with most lesions located in non-arterial structures such as veins (30%) or intrapulmonary opacifications (22%). There were on average 3.5 FP lesions (median 2.5, range 0–15) in the group of patients with PE and 5.0 (median 2, range 0–42) in the group without PE. In 72% of the examinations CAD indicated ≤5 FP, in 13% of the examinations CAD indicated ≥10 FP. Nine examinations had more than 20 FP lesions, and all of them were negative for the presence of PE according to the standard.

Discussion

The arterial pulmonary vasculature is a complex anatomical structure whose detailed analysis requires a structured reading approach for the detection of emboli. Thorough analysis of the pulmonary vasculature requires scrolling through the 300–500 axial CT sections multiple times, a time-consuming technique which requires continuous concentration.

The purpose of this study was to assess the performance of this CAD prototype and get an estimation of its potential diagnostic impact. We tested CAD as a standalone system. We chose to analyse CTPA examinations that had been performed during night shifts and weekends because they are often read by less experienced colleagues or under time pressure. For these reasons, these examinations may especially benefit from the availability of a CAD system. We compared the CAD results with our reference standard and the original reports.

Our results demonstrate that CAD picked up segmental and subsegmental emboli that were confirmed by the reference standard in seven patients who had been described not to have a PE in the original report. In two patients a solitary embolus had been missed in a segmental artery, in three patients a solitary subsegmental embolus was missed and in two patients two subsegmental emboli were missed.

Our results are compatible with previously published results from smaller studies that both included reader performance. In a study by Engelke et al., 56 CTPA were evaluated by two experienced and two inexperienced readers with

CAD as a second reader [17]. They showed the benefit of CAD for the detection of emboli at the segmental and subsegmental levels for the less experienced readers. Das et al. came to the same conclusions: at the subsegmental level the sensitivity of three readers (with 1 and 6 years of experience) increased with CAD from 80 to 92%, from 82 to 90% and from 63 to 81% [18]. However they only selected good quality images (vessel enhancement >200 HU) and excluded patients with underlying lung diseases. In both studies statistics were carried out only on a per-lesion basis.

It is important to note that in our study, we consecutively included all off-hour studies obtained within a certain time period. Thus we also explicitly included images of limited quality, e.g. on the basis of suboptimal enhancement or breathing artefacts. The study group also included images showing underlying lung disease, such as infiltrates or atelectases, another factor that may negatively influence the performance of a CAD system. In total we had to exclude 14 examinations: 8 because the CAD algorithm did not work for these data sets and 6 scans that were considered as non-diagnostic on the basis of massive streak artefacts.

Detection of emboli by CAD has been shown to depend on the level of the embolus (segmental or subsegmental) and the degree of obstruction. The study by Zhou et al. used 14 positive patients to test their CAD algorithm. Their sensitivity for partially (20–80%) occluded arteries was 84% for emboli in central, lobar and segmental arteries and dropped to 64% for emboli in subsegmental arteries [19]. If the vessel was minimally obstructed (≤20%), sensitivity dropped to 65 and 33% respectively. The sensitivity decreased further to 59 and 27% respectively with an obstruction of ≥80%. We included all obstruction levels and found the sensitivity of CAD as a standalone system to be 87% for the central, 78% for the lobar, 79% for the segmental and 61% for the subsegmental emboli.

On a per-patient basis, sensitivity of the CAD was 94%. Despite the variable image quality within the study group and the presence of underlying lung disease (n=200), CAD was able to identify 13 out of 14 patients with isolated subsegmental emboli. This emphasises the potential added value of the system for the detection of small emboli, a task for which human observers are also known to have reduced sensitivity. The clinical importance of isolated subsegmental emboli is still uncertain: it is suggested that isolated small emboli may be an indicator of deep venous thrombosis and therefore predictor of more severe embolic events in the future. Furthermore it was stated that the clinical relevance of small peripheral emboli is larger in individuals with cardiopulmonary restriction [8].

A large proportion (n=62/90, 69%) of subsegmental emboli in our study that

were missed by CAD were present in patients with central or lobar emboli. For these cases of central emboli, the detection of additional subsegmental emboli by CAD would have played a negligible clinical role unless the total embolus burden were used to direct patient management. Some subsegmental emboli (n=12/90, 13%) were missed because they were smaller than 2 mm, which is below the threshold of the CAD algorithm, or because they had led to complete obstruction with a sudden cessation of the subsegmental arteries, which was not detected by CAD.

CAD missed 4 patients out of 68 positive patients with in total one lobar embolus, four segmental emboli and one subsegmental embolus. These four false negative examinations were not characterised by especially low vascular enhancement or overlying motion artefacts. From a technical point of view, we were not able to identify obvious reasons why CAD was not able to identify the emboli in these four patients.

Most CTPA examinations done in clinical routine are eventually negative for the presence of PE. We found a negative predictive value (NPV) of 92% for examinations in which CAD did not find any candidate lesion. While an NPV of 92% is not yet sufficient for complete reliance on a negative CAD reading, it underlines the potential role of CAD as a second reading to reassure a reader when excluding PE. On the other hand, we have to state that only 44 out of the 210 negative cases did indeed have a completely negative CAD reading. Most negative PE examinations show differing numbers of false positive CAD lesions. The specificity of the algorithm on a per-patient basis is very low at 21% indicating that there is still a need for considerable improvement and that readers have to learn to efficiently rule out false positive candidate lesions if the application of CAD is to be beneficial.

Our results are comparable to those of the study by Walsham et al. [20]. They found an NPV of 84% and a specificity of 20% in a smaller (n=100), but also non-selected patient group. While Zhou et al. showed an average of 14.4 false positives per study, our system produced 4.7 false positives per study. Under the assumption that a maximum of 5 false positive lesions per examination marked by CAD represents an acceptable level in clinical routine, this criterion was met in 72% of our patients. Thus most examinations produce a relatively low number of false positive lesions. The median number of false positive lesions was 2 with most located in non-arterial structures such as veins (30%) or intrapulmonary opacifications (22%). Thus it is likely that in most scans the false positive lesions can easily be sorted out. The exact impact of the false positive lesions on diagnostic accuracy and reading time, however, remains to be determined in a reader study. More than 10 false positive lesions were seen in 13% of the CTPA examinations; not surprisingly, higher numbers of false

positive lesions correlated with lower image quality. Further reduction of false positive lesions per image is needed as too many may have a negative influence on diagnostic performance, lead to an unnecessarily prolonged reading time and may even cause unnecessary treatment.

Our study suffers from some limitations. There is no absolute reference standard for CTPA studies. Diagnostic pulmonary catheter angiography has been completely substituted by multidetector CTPA and had additionally been questioned as reference standard anyway. We therefore established a reference standard by independent evaluation of two readers with a third reader in case of discordant findings.

We evaluated the standalone performance of the CAD software to provide an estimate of the performance of the algorithm. In our study group, 10% of PE patients were originally missed by the on-call radiologist, all of whom had PE identified by CAD. However, comparison between the performance of CAD as a standalone system and that of the original reports is not sufficient to define the role of CAD in a clinical setting. The potential benefit of a CAD system strongly depends on the interaction with the reader, his or her experience and ability to discriminate between true- and false positive candidate lesions. More studies are necessary to evaluate the use of CAD as a second reading and to assess its impact on the reader's capability not only to detect emboli but also to exclude PE.

In summary, CAD can identify segmental and subsegmental pulmonary emboli that are missed by on-call radiologists, with an average of 4.7 false positive lesions per examination. A CTPA without any CAD candidates has a high negative predictive value, which may serve as reassurance for less experienced readers. Furthermore CAD found most of the isolated subsegmental emboli, which are clinically the most difficult emboli to find.

Reference List

1. British Thoracic Society. British Thoracic Society guidelines for the management of suspected acute pulmonary embolism. Thorax 2003; 58:470–483.
2. Olin JW. Pulmonary embolism. Rev Cardiovasc 2002; Med 3(Suppl 2):S68–S75.
3. Tapson VF. Acute pulmonary embolism. N Engl J Med 2008; 358:1037–1052.
4. Coche E, Pawlak S, Dechambre S et al. Peripheral pulmonary arteries: identification at multi-slice spiral CT with 3D reconstruction. Eur Radiol 2003; 13:815–822.
5. Ghaye B, Szapiro D, Mastora I et al. Peripheral pulmonary arteries: how far in the lung does multi-detector row spiral CT allow analysis? Radiology 2001; 219:629–636.
6. Remy-Jardin M, Remy J, Artaud D et al. Peripheral pulmonary arteries: optimization of the spiral CT acquisition protocol. Radiology 1997; 204:157–163.
7. Remy-Jardin M, Tillie-Leblond I, Szapiro D et al. CT angiography of pulmonary embolism in patients with underlying respiratory disease: impact of multislice CT on image quality and negative predictive value. Eur Radiol 2002; 12:1971–1978.
8. Schoepf UJ, Costello P. CT angiography for diagnosis of pulmonary embolism: state of the art. Radiology 2004; 230:329–337.
9. Schoepf UJ, Holzknecht N, Helmberger TK et al. Subsegmental pulmonary emboli: improved detection with thin-collimation multi-detector row spiral CT. Radiology 2002; 222:483–490.
10. Torbicki A, Perrier A, Konstantinides S et al. Guidelines on the diagnosis and management of acute pulmonary embolism: the task force for the diagnosis and management of acute pulmonary embolism of the European Society of Cardiology (ESC). Eur Heart J 2008; 29:2276–2315.
11. van Belle A, Buller HR, Huisman MV et al. Effectiveness of managing suspected pulmonary embolism using an algorithm combining clinical probability, D-dimer testing, and computed tomography. JAMA 2006; 295:172–179.
12. Chan HP, Hadjiiski L, Zhou C et al. Computer-aided diagnosis of lung cancer and pulmonary embolism in computed tomography - a review. Acad Radiol 2008; 15:535–555.
13. Schaefer-Prokop C, Prokop M. MDCT for the diagnosis of acute pulmonary embolism. Eur Radiol 2005; 15 (Suppl 4):D37–D41.
14. Boyden E. Segmental anatomy of the lungs. New York, NY: McGraw-Hill, 1955.
15. Jackson C, Huber J. Correlated applied anatomy of the bronchial tree and lungs with a system of nomenclature. Dis Chest 1943; 9:319–326.
16. Buelow T, Wiemker R, Blaffert T et al. Automatic extraction of the pulmonary artery tree from multi-slice CT data. Proc SPIE 2005; 5746:730–740.
17. Engelke C, Schmidt S, Bakai A et al. Computer-assisted detection of pulmonary embolism: performance evaluation in consensus with experienced and inexperienced chest radiologists. Eur Radiol 2008; 18:298–307.
18. Das M, Muhlenbruch G, Helm A et al. Computer-aided detection of pulmonary embolism: influence on radiologists' detection performance with respect to vessel segments. Eur Radiol 2008; 18:1350–1355.

19. Zhou C, Chan HP, Patel S et al. Preliminary investigation of computer aided detection of pulmonary embolism in three-dimensional computed tomography pulmonary angiography images. Acad Radiol 2005; 12:782–792.
20. Walsham AC, Roberts HC, Kashani HM et al. The use of computer aided detection for the assessment of pulmonary arterial filling defects at computed tomographic angiography. J Comput Assist Tomogr 2008; 32:913–918.

Chapter 4

Wittenberg R
Peters JF
Sonnemans JJ
Bipat S
Prokop M
Schaefer-Prokop CM

Am J Roentgenol. 2011 Jan; 196(1):95-101.

Impact of image quality on the performance of a computer-assisted detection prototype for pulmonary embolism

Abstract

Purpose

The purpose of this article is to assess the relationship between CT image quality and the number and type of false positive (FP) findings found by a prototype computer-assisted detection (CAD) algorithm for automatic detection of pulmonary embolism (PE).

Materials and methods

This retrospective study included 278 subjects (138 men and 140 women; mean age 57 years; range 18–88 years) who underwent consecutive CT pulmonary angiographies performed during off hours. Twenty-four percent (68/278) of studies were reported as positive for PE. CAD findings were classified as true positive or FP by two independent readers and, in cases of discordance, by a third radiologist. Each FP result was classified according to the underlying cause. The degree of vascular enhancement, image noise, motion artefacts, overall quality, and presence of underlying lung disease were rated on a 4- or 5-point scale. Chi-square tests and *t*-tests were used to test significance of differences.

Results

The mean number of FP CAD findings was 4.7 (median 2) per examination. Most were caused by veins (30% (389/1298)) or airspace consolidations (22% (286/1298)). There was a significant positive association between the number of FP findings and image noise, motion artefacts, low vascular enhancement, low overall quality, and the extent of underlying disease. On a per-embolism basis, sensitivity decreased from 70.6% (214/303) for scans with zero to five FP findings, to 62.3% (33/53) for scans with six to 10 FP findings and to 60% (12/20) for scans with more than 10 FP findings.

Conclusions

There is a strong association between CT image quality and the number of FP findings indicated by a CAD algorithm for the detection of PE.

Introduction

Acute pulmonary embolism (PE) is a potentially life-threatening disease and is a common problem in emergency units, especially during night or weekend shifts. Multidetector computed tomography (MDCT) has been established as the method of choice for diagnosing acute PE [1]. It has a significantly higher sensitivity for diagnosis of small peripheral emboli than single-detector CT does, but it produces approximately two to three times more axial sections, which have to be scrutinized for the presence of PE [2-5].

The use of computer-assisted detection (CAD) tools for the detection of PE is new. It is known from other diagnostic tasks (e.g., detection of pulmonary nodules in CT) that a multitude of images to be evaluated and complex overlying anatomic structures represent distracting features that may lower the detection performance of the observer [6,7]. Tracing of the entire pulmonary artery tree down to the subsegmental arteries to detect potential filling defects is also time consuming. At the same time, the majority of scans are made to exclude PE, because the prevalence of PE in patients in whom the disease is suspected is low, ranging from 5% to 35%, depending on patient selection and referral practices [8-10]. Thus, exclusion of PE is rather time consuming and small segmental and subsegmental emboli may still be missed.

CAD algorithms for PE have been developed to alert radiologists to suspicious lesions to speed up the evaluation process and to reduce the number of false negative (FN) diagnoses. Although most false positive (FP) CAD findings can be easily identified by the reader, a high number of FP candidates per scan increases interpretation time and can increase the number of FP readings. Studies thus far mainly have focused on the influence of various CAD algorithms on the performance of the radiologist [11-14] or have tested the standalone performance of a CAD algorithm [15-20]. Results will not necessarily be the same in a group of consecutive pulmonary CT angiography examinations that also includes low-quality examinations as compared with a group of scans selected for optimal image quality. An estimation of the CAD algorithm's performance under less favourable conditions, however, is important before applying such an algorithm in clinical routine. To date, with the exception of two studies that both included rather small study groups [19,20], to our knowledge none of the previous studies has attempted to quantify the relationship between image quality and the number of FP findings made with CAD software.

The purpose of this study was to assess the relationship between CT image quality and the number and type of FP lesions found by a CAD algorithm for automatic detection of PE.

Materials and Methods

Patient selection

This retrospective data analysis was approved by the local ethics committee. The study included all 292 pulmonary CT angiographies that were consecutively acquired during night and weekend shifts between January 2007 and April 2008 in a university hospital. Six examinations were excluded because of massive streak artefacts either caused by arms placed parallel to the chest or because of metallic material that made diagnostic evaluation of these scans impossible. Eight examinations had to be excluded because the CAD algorithm failed: in one case, the patient had a left-sided pneumothorax and in another case, the left lung had been resected. The other six patients had a tracheal tube for ventilation, which caused a connection between extrathoracic air and the lungs and subsequent failure of segmentation of the trachea and lungs.

Thus, the final study group consisted of 278 examinations in 278 patients: 138 men and 140 women with a mean age of 57 years (range, 18–88 years). There were five patients from the outpatient clinic, 140 patients admitted to the emergency unit, and 133 inpatients (13 of whom came from the ICU).

CT technique

Examinations of 178 patients were performed on a 16-MDCT scanner (MX 8000 IXDT or Brilliance-16, Philips Healthcare); 100 patients were scanned on a 64-MDCT scanner (Brilliance-64, Philips Healthcare). All scans were obtained in a caudocranial direction from the level of the diaphragm to the lung apices within a single breath-hold.

The 64-slice scans were obtained with 64 × 0.625 mm collimation, a rotation time of 0.42 second, and a pitch of 1.172. Exposure parameters were 120 kV and 100 mAs. All images were reconstructed with 0.9-mm slice thickness and 0.45-mm overlap. The 16-slice scans were obtained with 16 × 0.75 mm collimation, a rotation time of 0.5 second, and a pitch of 1.188. Exposure parameters were 90 kV and 180 mAs. All images were reconstructed with 1.0-mm slice thickness and 0.5-mm overlap.

A dual-head CT injector (Stellant, Medrad Europe) was used for IV injection of contrast medium. All patients received 90 ml of contrast medium IV (iopromide; Ultravist 300, Bayer Schering Pharma). Contrast medium was injected with a flow rate of 5.0 ml/s on the 64-MDCT scanner, followed by a 30-ml saline chaser bolus. On the 16-MDCT scanner, contrast was injected with a flow rate of 4.0 ml/s, followed by 40-ml saline chaser bolus. All examinations were started by bolus triggering in the pulmonary trunk. The trigger threshold was set at

150 HU, and a start delay of 8 seconds after reaching the trigger threshold was used.

CAD algorithm

All examinations were analysed using a prototype CAD algorithm (Philips Healthcare). This algorithm consists of the following steps: lung and airway segmentation, vessel segmentation, analysis of contrast differences in the vessels, candidate generation, and filtering to reduce FP findings.

First, the lung and airways are segmented using a combination of 2D and 3D image-processing techniques. Second, a vessel-tracking algorithm is applied to extract the pulmonary vessels, starting with segmentation of all structures above –100 HU [21,22]. For each branch of the pulmonary vascular tree, a centreline and ring following the vessel wall are created. The algorithm looks for local contrast differences within the vessels beyond a certain threshold. These contrast variations are analysed and the obstruction is segmented. From partial or complete obstruction segmentations, candidate lesions are generated. Finally, features of the candidate lesions, such as shape, size, and Hounsfield unit statistics are computed to filter FP lesions using a decision tree. The remaining candidate lesions found by the CAD algorithm are surrounded by circles and presented to readers on demand. This analysis takes approximately 30 seconds on a 2.1-GHz processor (Xeon, Intel) and runs in the background during initial reading.

Data collection and analysis

A reference for the presence of PE was established by two independent readers: a researcher specially trained in reading PE studies (> 250 scans) and a chest radiologist with more than 15 years of experience. Discordant findings between the two readers or between the readers and the original report were determined by another chest radiologist with more than 15 years of experience. The readers were asked to digitally mark all filling defects within the contrast-enhanced lumen of the pulmonary arteries separately. The anatomic level of the thrombus was annotated (central, lobar, segmental, and subsegmental) using the nomenclature derived from Boyden [23] and Jackson and Huber [24]. The CAD findings were compared with the reference standard and were scored as true positive (TP), FP, or FN (Figure 1 and 2).

All FP findings were classified with regard to the most likely underlying reason for the misinterpretation. Underlying anatomic reasons for the FP interpretation included pulmonary veins (inhomogeneously opacified), lymphoid tissue, airways (mucus filled), airspace consolidation or arterial branching. Reasons for FP findings related to scanning technique were motion artefacts or inadequate arterial enhancement.

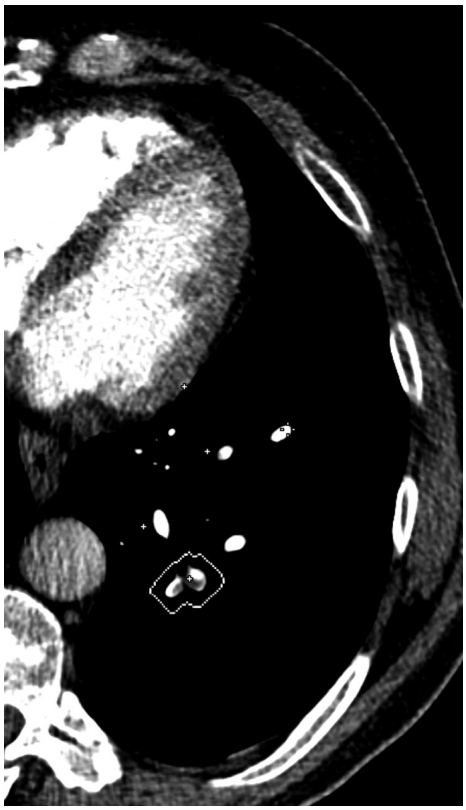

Figure 1. Small pulmonary embolus found by a computer-assisted detection algorithm.

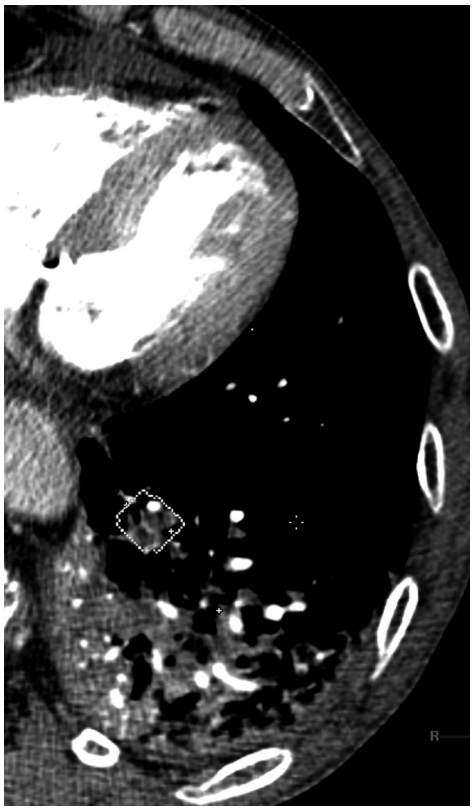

Figure 2. False positive computer-assisted detection finding due to accompanying lung disease.

The quality of the CT examinations was subjectively scored without knowledge of the CAD findings using a 5-point scale with respect to overall quality, vascular enhancement, image noise, and motion artefacts. A score of 1 represented inadequate image quality that made diagnosis impossible and a score of 5 represented excellent image quality. The effect of accompanying lung diseases on diagnosis was classified using a 4-point scale that rated the effect of parenchymal disease on the detection of PE by the radiologists or the CAD algorithm, respectively (Table 1). Image noise was also measured by the researcher as the SD of CT numbers in a circular region of interest of 1 cm^2 in the descending aorta at the level of the bifurcation of the main pulmonary arteries. Measurements performed on the 1-mm sections were repeated three times and subsequently averaged.

Table 1. Subjective scoring of image quality and grading of accompanying lung disease

Parameter, Score	Description of Score
Overall image quality or arterial enhancement	
1	Inadequate, no diagnosis of PE[A] possible
2	Low, diagnosis possible until the lobar level
3	Sufficient, diagnosis possible until the segmental level
4	Good, diagnosis possible until the subsegmental level
5	Excellent
Noise or motion artefacts	
1	Massive, no diagnosis possible
2	Definite, establishment of diagnosis impeded
3	Moderate, but image sufficient for diagnosis of PE
4	Minor
5	None
Accompanying lung disease	
1	Present, lung disease disturbing for CAD[B] and radiologist
2	Present, lung disease disturbing for CAD
3	Present, but no influence
4	None

[A] Pulmonary embolism.
[B] Computer-assisted detection

Statistical analysis

A statistical analysis was performed with SPSS (version 15.0, SPSS). To assess the relationship between CAD performance and parameters of image quality, the following statistical analyses were performed:

A univariate analysis was done using the number of FP results as the dependent outcome variable and overall quality, vascular enhancement, noise, motion artefacts, and accompanying lung diseases as independent variables. Variables were considered explanatory if the p-values were less than 0.05 using the chi-square test.

A univariate logistic regression analysis was done to predict TP and FP CAD lesions as a function of image quality. TP and FP CAD lesions were used as dependent variables and overall quality, vascular enhancement, noise, motion

artefacts and accompanying lung disease were used as independent variables. Subsequently, covariates with a p-value of less than 0.05 were included in the multivariate logistic regression model using stepwise backward selection.

A Spearman's rank's correlation coefficient was used to correlate the measured noise and the subjective noise score. The Mann-Whitney U test was used to test for significance of differences between the two subgroups of examinations obtained with the 16- and 64-MDCT scanners. Additionally, we calculated the standalone sensitivity of the CAD algorithm on a per-patient basis and on a per-lesion basis.

Results

Reference standard

Of the 278 scans, our reference standard found 68 (24%) positive scans. For a total of 63 lesions in 42 patients, a third experienced chest radiologist was consulted as an arbiter. Fifty-five of the 63 lesions were located in subsegmental arteries and the remaining eight were located in segmental arteries. These discordant findings were present in 36 patients with PE and six patients without PE with a mean quality score of 4.0 (median 4; range 2–5).

FP findings analysis

CAD found a total of 258 TP findings and 1298 FP findings (n = 1556). The mean number of FP findings marked by the CAD algorithm was 4.7 per examination (median 2; range 0–42). The mean number of FP findings was 3.5 (median 2.5; range 0–15) in the group of patients with PE and 5.0 (median 2; range 0–42) in the group without PE. In 72% (200/278) of the examinations, CAD indicated five or fewer FP findings and in 13% (35/278) of the examinations, CAD indicated more than 10 FP findings.

Most FP findings were located in veins (30% (389/1298)) or in areas with airspace consolidations (22% (286/1298)). The remaining FP findings were caused by motion artefacts (16% (208/1298)), inadequate arterial enhancement (14% (182/1298)), lymphoid tissue (13% (169/1298)), arterial branching (3% (39/1298)) and mucus-filled airways (2% (26/1298)) (Figure 3).

Impact of overall image quality

The mean overall quality score was 4.0 (median 4; range 1–5). The overall quality was graded as excellent (score 5) in 31.3% (87/278) of cases and as good (score 4) in 44.6% (124/278) of cases. In only 0.7% (2/278) of the cases, the overall quality was judged as inadequate (score 1) (Table 2).

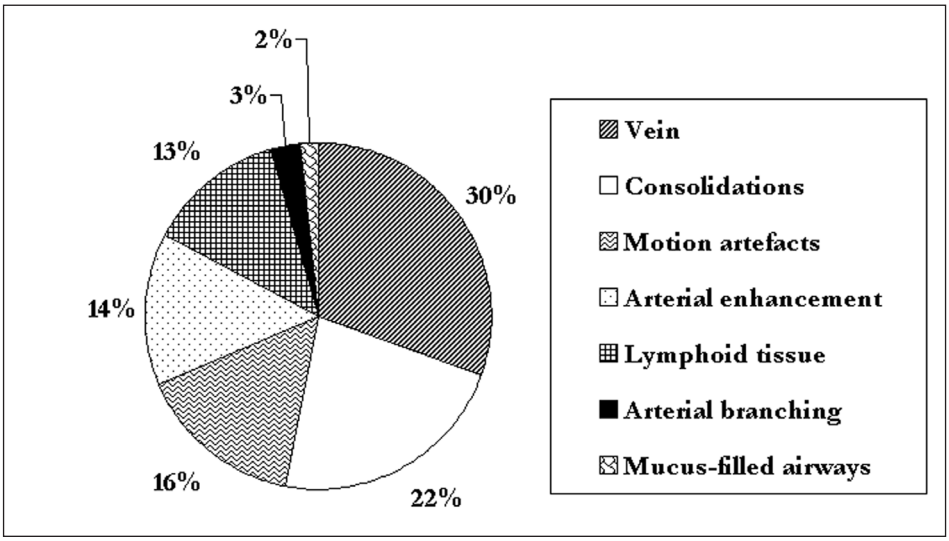

Figure 3. Classification of underlying reasons for false positive computer-assisted detection findings.

In the group of scans that were judged to allow evaluation of the complete vascular tree down to the subsegmental level, five or fewer FP findings were seen in 94% (82/87) of cases with a score of 5 and in 77% (95/124) of cases with a score of 4. In scans that were judged to allow evaluation only down to the segmental and lobar level, five or fewer FP findings were seen in only 45% (21/47) of cases with a score of 3 and in 22% (4/18) of cases with a score of 2. None of the inadequate scans (score 1) had five or fewer FP findings. The mean number of veins marked by CAD continuously decreased from 5 to 3.2, 3.1, 1.2, and 0.4, respectively, with increasing overall image quality scores.

Impact of vascular enhancement

The mean vascular enhancement score was rated as 4.5 (median 5; range 1–5). In 68.7% (191/278) of the cases, the vascular enhancement was scored as excellent (score 5) and in 16.9% (47/278) it was scored as good (score 4). In only one case, low enhancement precluded diagnosis (Table 2).

In the group of scans that were judged to allow evaluation of the complete vascular tree down to the subsegmental level, five or fewer FP findings were seen in 84% (161/191) of cases with a score of 5 and in 66% (31/47) of cases with a score of 4. In scans that were judged to allow evaluation only down to the segmental and lobar level, five or fewer FP findings were seen in only 43% (10/23) of cases with a score of 3 and in 13% (2/16) of cases with a score of 2. The inadequate scans (score 1) had more than five FP findings.

Table 2. Frequency (Freq) of scans per image quality score and mean number of false-positive (FP) findings per scan for 278 patients.

	Overall Quality		Vascular Enhancement		Noise		Motion Artefacts		Accompanying Lung Disease	
Score	Freq (%)	FP Findings / Scan	Freq (%)	FP Findings / Scan	Freq (%)	FP Findings / Scan	Freq (%)	FP Findings / Scan	Freq (%)	FP Findings / Scan
1	0.7	24.5	0.4	7	0.4	7	1.4	8.8	2.1	9.8
2	6.5	11.8	5.8	14.8	6.8	15.2	11.8	8	35.6	4.8
3	16.9	9.1	8.3	8.8	53.6	5	57.6	4.9	34.2	3.2
4	44.6	3.8	16.9	5.4	33.1	2.5	28.1	2.7	28.1	2.4
5	31.3	1.6	68.7	3.1	6.1	1.5	1.1	2	NA	NA

Note. Score of 1 refers to inadequate overall quality and vascular enhancement, high image noise and disturbing motion artefacts. Score of 5 refers to excellent overall quality and vascular enhancement and the absence of noise and motion artefacts. The presence of accompanying lung diseases was scored using a 4-point scale, with 1 referring to disturbing accompanying lung diseases and 4 to the absence of accompanying lung diseases (see Table 1). NA = not applicable .

Impact of motion artefacts

The mean score for the presence of motion artefacts was rated as 3.2 (median 3; range 1–5) for the total group. Except for 1.1% (3/278) of the cases (score 5), all scans showed variable degrees of motion artefacts. In 28.1% (78/278) of the cases, the presence of motion artefacts was rated as minor (score 4). In the majority of scans (57.6% (160/278)) the presence of motion artefacts was moderate, but the image was still sufficient for diagnosis of PE (score 3). In only 1.4% (4/278) of the cases, the presence of motion artefacts was massive and no diagnosis of PE was possible (Table 2). Five or fewer FP findings were seen in 100% (3/3) of the scans with no motion artefacts (score 5) and 85% (66/78) of the scans with minor motion artefacts (score 4).

Impact of image noise

The mean score for the image noise was 3.4 (median 3; range 1–5). Except for 6.1% (17/278) of the cases (score 5), noise was present in variable degrees in all other scans. In 33.1% (92/278) of the cases, the presence of noise was rated as minor (score 4). In the majority of scans (53.6% (149/278)) the presence of noise was moderate, but the image was still sufficient for diagnosis of PE (score 3). In 0.4% of the scans (n=1), the presence of noise was massive precluding the diagnosis of PE (Table 2). Five or fewer FP findings were seen in 94% (16/17) of the scans with no noise (score 5) and in 88% (81/92) of the scans with minor noise (score 4).

The mean image noise measured in the descending aorta at the level of the bifurcation of the main pulmonary arteries was 32 HU (median 29 HU; range 13–84 HU). There was a significant correlation ($p<0.001$) between the measured noise and the subjective noise score using the Spearman's rank correlation coefficient.

Impact of accompanying lung diseases

In 28.1% (78/278) of the cases, there was no accompanying disease (score 4). The mean number of FP findings in this group was 2.4 and the mean overall quality score was 4.2. In 34.2% (95/278) of the cases, accompanying lung disease was present but not disturbing for the CAD software or the radiologist (score 3). The mean number of FP findings was 3.2 and the mean overall quality score was 4.1 in this patient group. In 35.6% (99/278) of the cases, the CAD software marked a contrast difference caused by airspace consolidations as PE although the radiologist did not feel hampered in his diagnosis by the disease (score 2). The mean number of FP findings was 4.8 in this patient group and the mean overall quality score was 3.8. In 2.1% (6/278) of these cases, the accompanying disease was perceived also by the radiologist as disturbing for establishing the

diagnosis of PE (score 1). The mean number of FP findings was 9.8 and the mean overall quality score was 3.0 in these patients.

Impact of scanner type

One hundred seventy-eight examinations were performed with a 16-MDCT scanner and 100 examinations were performed with a 64-MDCT scanner. Examinations performed with the 16-MDCT scanner showed a lower image noise (34.3 vs 27.7 HU; p<0.001), higher subjective scores of vascular enhancement (4.7 vs 4.1; p<0.001) and higher subjective scores of overall quality (4.1 vs 3.8; p=0.001). There was no difference with respect to the impact of accompanying lung diseases (2.9 vs 2.8; p=0.5) and the severity of motion artefacts (3.4 vs 3.3; p=0.6).

Examinations performed with the 16-MDCT scanner showed a mean number of 3.8 FP findings per scan (median 2; range 0–42), compared with a mean of 6.3 FP findings per scan (median 4.5; range 0–36) in examinations performed with the 64-MDCT scanner, with the difference being significant (p=0.001).

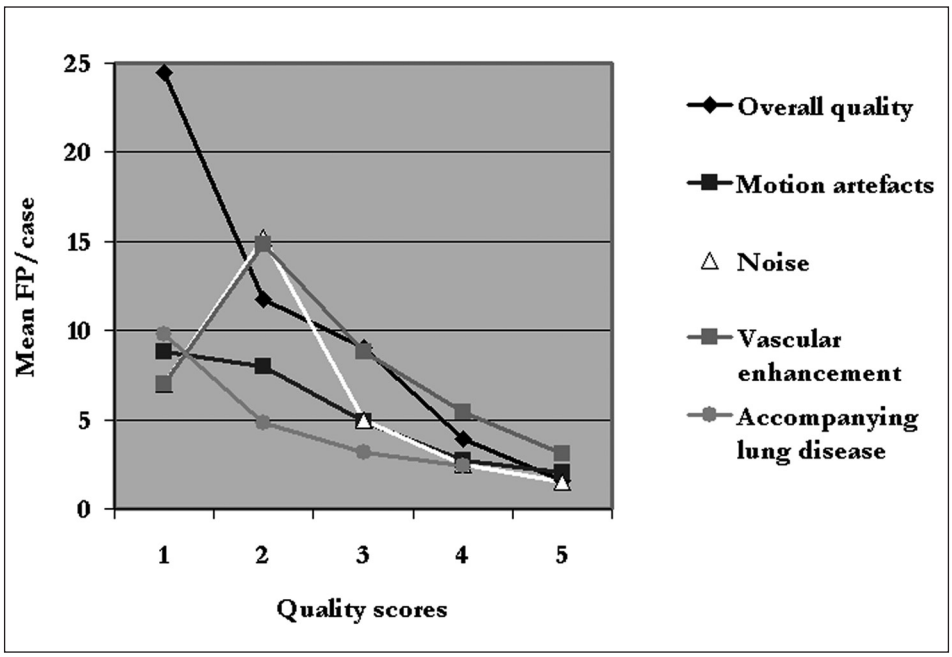

Figure 4. Significant association between mean number of false positive (FP) findings per case and subjectively measured overall quality, motion artefacts, noise, vascular enhancement and accompanying lung disease.

Predictive values

A chi-square test showed significant associations between the number of FP findings and overall quality, vascular enhancement, motion artefacts, noise and accompanying lung diseases ($p<0.001$) (Table 4). A multivariate analysis using logistic regression showed that the overall quality ($p<0.001$; odds ratio 4.5), vascular enhancement ($p=0.025$; odds ratio 0.74), noise ($p=0.002$; odds ratio 0.68) and accompanying lung diseases ($p<0.001$; odds ratio 2.2) are predictors of TP and FP CAD lesions. Only for the presence of motion artefacts could no significant ($p=0.88$) relationship be shown in this multivariate model.

Sensitivity

The overall sensitivity of the CAD was 94% (64/68) on a per-patient basis. On a per-lesion basis, the sensitivity was 70.6% (214/303) for the subgroups of scans with up to five FP findings, 62.3% (33/53) for the scans with six to 10 FP findings and 60% (12/20) for the scans with more than 10 FP findings.

Analysis of negative PE studies

The overall specificity of the CAD was 21% (44/210) on a per-patient basis. Ninety-two percent (44/48) of the examinations without any candidate lesion were also judged to be negative for PE by the reference standard. The mean quality score for all negative PE studies was 4.5 for the overall quality, 4.9 for the vascular enhancement, 3.5 for the motion artefacts, 3.7 for the noise, and 3.7 for accompanying lung diseases.

Fifty-nine percent (26/44) of the negative studies with no FP findings had excellent overall quality (score 5), 36% (16/44) had a good quality (score 4), and 5% (2/44) had a moderate quality (score 3). None of the negative studies with no FP findings had a low or inadequate quality (score 1 and 2).

Discussion

CAD algorithms have been developed with the objectives of helping exclude PE more reliably, improving the detection by radiologists and, ideally, speeding up the evaluation. Whether and to what degree these options can be fulfilled by CAD is currently under evaluation. It is very likely that the quality of pulmonary CT angiography influences the number of FP findings found by CAD algorithms. Poorly timed contrast injection, motion artefacts, and noise all directly influence the intravascular contrast homogeneity and, thus, the detection of a CAD algorithm that looks for intravascular contrast differences. Also, accompanying lung disease may reduce the performance of the CAD algorithm,

similar to how lung disease makes evaluation more difficult for radiologists [25].

The purpose of this article was therefore to assess the relationship between CT image quality and the number and type of FP findings found by a CAD prototype for automatic detection of PE. To get a realistic estimate of the performance of CAD under clinical conditions, all pulmonary CT angiography examinations performed over a period of 16 months during off hours were consecutively included, irrespective of image quality. For these reasons, these examinations may especially benefit from the availability of a CAD system.

For the CAD prototype evaluated in this study, we found a strong association between the number of FP CAD findings and the overall quality, motion artefacts, vascular enhancement, noise and accompanying lung diseases. Furthermore, a multivariate logistic regression analysis showed that overall quality, vascular enhancement, noise and accompanying lung diseases can predict whether CAD lesions are TP or FP. The presence of motion artefacts was not significant in this model, because it was highly correlated with the other image quality parameters.

We found a mean of 4.7 FP findings per scan, with one third of the FP candidates located in pulmonary veins. The number of FP findings was higher in examinations performed with the 64-MDCT scanner, most likely because of the higher image noise and lower vascular enhancement of those examinations, as compared with the examinations performed with the 16-MDCT scanner. Spatial resolution (0.9- and 1.0-mm slice thickness) and the presence of motions artefacts were comparable. These findings further underline the association between image quality parameters, especially image noise and vascular enhancement and the specificity of the CAD algorithm. Both parameters are dependent on the protocol chosen and not on the scanner type. We therefore do not think that there is an association between scanner type and the performance of the CAD algorithm and do not conclude a general superiority of 16-MDCT scanners.

Our results are comparable with the findings reported in other studies. Das et al. [11] found on average four FP findings per scan, with most of them caused by misinterpretation of venous fillings, the azygos vein and lymphoid tissue. Similar to Das et al. we found veins (30% (389/1298)), airspace consolidation (22% (286/1298)) and lymphoid tissue (13% (169/1298)) as the most important reasons for misinterpretations of contrast differences as FP findings. We found a decreasing number of FP CAD findings located in veins with increasing overall quality. This can be explained by the fact that we had classified any candidate lesion located in a vein as FP irrespective of whether inhomogeneous venous enhancement, airspace consolidations, or motion artefacts contributed to the FP interpretation of the CAD algorithm. Thus, the decreasing number of FP findings in veins with increasing image quality does not reflect improved anatomic differentiation but other aspects of image quality, such as decreasing

presence of motion artefacts or increased intravascular contrast.

We concluded from these results that an increased CAD performance can most effectively be achieved by making arterial-venous separation an integral part of the algorithm. Adapting the contrast injection protocol in a way that both arteries and veins are more homogeneously contrast enhanced may be an alternative option to at least decrease the number of FP findings, even though reasons such as motion artefacts or anatomic misinterpretation pertain.

Our study group consisted of consecutive pulmonary CT angiography examinations that were not selected with respect to image quality. To get an estimate of how many of these examinations to which CAD could be potentially applied as a second reader, we made a special effort to assess the number and quality of examinations that had five or fewer FP findings. This cutoff was chosen in accordance with CAD algorithms applied in chest radiography that also use an upper limit of five candidate lesions. Second, we considered this a number of candidates that appears to be acceptable for routine application. Looking at the results, 72% (200/278) of the examinations had five or fewer FP findings and all of these scans showed excellent to good assessments (scores 5 and 4) for the various parameters. We concluded from this finding that, in addition to further improvement of the CAD algorithm itself, good image quality seems to be a crucial point if CAD is aimed to be applied effectively under clinical conditions. Twelve percent (33/278) of examinations had more than 10 FP findings, a fact that certainly precludes a meaningful application of CAD. This is especially noteworthy because these examinations represent the group of CT pulmonary angiography scans that are also more challenging for less experienced readers.

Though Das et al. [11] only analysed scans with good image quality (vessel enhancement > 200 HU) and explicitly excluded patients with underlying lung diseases, we found a comparable number of FP findings (4.7 per scan). We found a strong association between vascular enhancement and the number of FP findings. Because our study group also included scans with suboptimum enhancement and breathing artefacts, it is likely that the absolute number of FP findings per underlying reason was different from that in the study by Das et al., though the relative distribution seemed similar.

Zhou et al. [20] developed their own CAD algorithm for the detection of PE. They compared six positive PE scans without accompanying lung diseases with eight positive scans with lung diseases. They showed that additional lung disease has a significant influence on the performance of CAD, because the sensitivity to detect PE decreased from 92% to 66.7% for central, lobar, and subsegmental arteries and from 77.8% to 40% for subsegmental arteries. We also found a strong association between the number of FP findings and the presence of accompanying lung disease.

We observed that, on a per-lesion basis, the sensitivity decreased from 70.6% (214/303) for the subgroup of scans with zero to five FP findings, to 62.3% (33/53) for the scans with six to 10 FP findings, to 60% (12/20) for the scans with more than 10 FP findings. The less pronounced differences in sensitivity suggest that the sensitivity of the CAD algorithm is less susceptible to image quality than the number of FP findings. Similar results were recently reported by Dewailly et al. [26], who found no significant correlation between image quality and the sensitivity of a CAD prototype different from the one we tested. We did not attempt to estimate the direct effect of image quality on the standalone sensitivity of the CAD, because the number of examinations with moderate-to-poor overall quality was substantially smaller than the number of examinations with good or excellent quality. Furthermore, the number of examinations found to be positive for PE in this small group was insufficient for reliable statistics.

Our study has some limitations. There is no absolute standard of truth for the presence of PE in pulmonary CT angiography studies. Our reference standard was established by a researcher and an experienced chest radiologist, with a third experienced radiologist in cases of discordant findings. This reference standard is suboptimal, but it reflects clinical reality, where CT has become the standard of reference for PE diagnosis. Though a non-radiologist, but well-trained researcher, took part in the reading process, it has to be noted that all studies were seen by a radiologist during clinical evaluation and an experienced chest radiologist during study evaluation. Furthermore, all discrepancies were seen by a third radiologist.

It has been previously shown that interobserver agreement decreases with lower image quality [27, 28]. Thus, especially in pulmonary CT angiography studies of lower image quality in which CAD performance decreased, the reference also has to be considered potentially weaker. However, given the fact that those 42 examinations in which discrepant findings had required the consultation of the arbiter did not show a lower image quality score, we consider the effect rather small.

In our study, we analysed consecutively obtained pulmonary CT angiography examinations that all were performed during night shifts and weekends in a single institution. Different scan protocols, scanner types, or composition of study group may yield different results, although it is likely that the positive association between image quality and CAD performance will persist.

In summary, we have found a strong association between image quality of pulmonary CT angiography scans and the number of FP findings indicated by the specific CAD prototype tested. The mean number of FP findings was 4.7, with most findings located in veins. Further improvement of the software is needed to make the performance of these algorithms less vulnerable to image

artefacts caused by motion, low vascular contrast and airspace consolidations. The specificity of the prototype could be substantially improved if the CAD algorithm included a differentiation between veins and arteries. Alternatively, an adaptation of the contrast application protocol achieving high arterial and venous intravascular contrast may help to reduce the number of FP findings located in veins.

Reference List

1. Quiroz R, Kucher N, Zou KH et al. Clinical validity of a negative computed tomography scan in patients with suspected pulmonary embolism: a systematic review. JAMA 2005; 293:2012-2017.
2. Remy-Jardin M, Pistolesi M, Goodman LR et al. Management of suspected acute pulmonary embolism in the era of CT angiography: a statement from the Fleischner Society. Radiology 2007; 245:315 -329.
3. Stein PD, Fowler SE, Goodman LR et al. Multidetector computed tomography for acute pulmonary embolism. N Engl J Med 2006; 354:2317 -2327.
4. Schoepf UJ, Holzknecht N, Helmberger TK et al. Subsegmental pulmonary emboli: improved detection with thin-collimation multi-detector row spiral CT. Radiology 2002; 222:483-490.
5. Winer-Muram HT, Rydberg J, Johnson MS et al. Suspected acute pulmonary embolism: evaluation with multi-detector row CT versus digital subtraction pulmonary arteriography. Radiology 2004; 233:806 -815.
6. Petrick N, Haider M, Summers RM et al. CT colonography with computer-aided detection as a second reader: observer performance study. Radiology 2008; 246:148-156.
7. White CS, Pugatch R, Koonce T et al. Lung nodule CAD software as a second reader: a multicentre study. Acad Radiol 2008; 15:326 -333.
8. Torbicki A, Perrier A, Konstantinides S et al. Guidelines on the diagnosis and management of acute pulmonary embolism: the Task Force for the Diagnosis and Management of Acute Pulmonary Embolism of the European Society of Cardiology (ESC). Eur Heart J 2008; 29:2276-2315.
9. van Belle A, Buller HR, Huisman MV et al; Christopher Study Investigators. Effectiveness of managing suspected pulmonary embolism using an algorithm combining clinical probability, D-dimer testing, and computed tomography. JAMA 2006; 295:172-179.
10. Herédia V, Ramalho M, Zapparoli M et al. Incidence of pulmonary embolism and other chest findings in younger patients using multidetector computed tomography. Acta Radiol 2010; 51:402 -406.
11. Das M, Muhlenbruch G, Helm A et al. Computer-aided detection of pulmonary embolism: influence on radiologists‘ detection performance with respect to vessel segments. Eur Radiol 2008; 18:1350 -1355.
12. Engelke C, Schmidt S, Bakai A et al. Computer-assisted detection of pulmonary embolism: performance evaluation in consensus with experienced and inexperienced chest radiologists. Eur Radiol 2008; 18:298 -307.
13. Walsham AC, Roberts HC, Kashani HM et al. The use of computer-aided detection for the assessment of pulmonary arterial filling defects at computed tomographic angiography. J Comput Assist Tomogr 2008; 32:913-918.
14. Engelke C, Schmidt S, Auer F et al. Does computer-assisted detection of pulmonary emboli enhance severity assessment and risk stratification in acute pulmonary embolism? Clin Radiol 2010; 65:137 -144.

15. 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.
16. Maizlin ZV, Vos PM, Godoy MC et al. Computer-aided detection of pulmonary embolism on CT angiography: initial experience. J Thorac Imaging 2007; 22:324-329.
17. Masutani Y, MacMahon H, Doi K. Computerized detection of pulmonary embolism in spiral CT angiography based on volumetric image analysis. IEEE Trans Med Imaging 2002; 21:1517-1523.
18. Schoepf UJ, Schneider AC, Das M et al. Pulmonary embolism: computer-aided detection at multidetector row spiral computed tomography. J Thorac Imaging 2007; 22:319 -323.
19. Wittenberg R, Peters JF, Sonnemans JJ et al. Computer-assisted detection of pulmonary embolism: evaluation of pulmonary CT angiograms performed in an oncall setting. Eur Radiol 2010; 20:801 -806.
20. Zhou C, Chan HP, Patel S et al. Preliminary investigation of computer-aided detection of pulmonary embolism in three-dimensional computed tomography pulmonary angiography images. Acad Radiol 2005; 12:782 -792.
21. Bouma H, Sonnemans JJ, Vilanova A et al. Automatic detection of pulmonary embolism in CTA images. IEEE Trans Med Imaging 2009; 28:1223 -1230.
22. Buelow T, Wiemker R, Blaffert T et al. Automatic extraction of the pulmonary artery tree from multislice CT data. Proc SPIE 2005; 5746:730 -740.
23. Boyden E. Segmental anatomy of the lungs. New York, NY: McGraw-Hill, 1955.
24. Jackson C, Huber J. Correlated applied anatomy of the bronchial tree and lungs with a system of nomenclature. Dis Chest 1943; 9:319 -326.
25. Schoepf UJ, Costello P. CT angiography for diagnosis of pulmonary embolism: state of the art. Radiology 2004 ; 230:329 -337.
26. Dewailly M, Remy-Jardin M, Duhamel A et al. Computer-assisted detection of acute pulmonary embolism with 64-slice multi-detector row computed tomography: impact of the scanning conditions and overall image quality in the detection of peripheral clots. J Comput Assist Tomogr 2010; 34:23 -30.
27. Ginsberg MS, King V, Panicek DM. Comparison of interpretations of CT angiograms in the evaluation of suspected pulmonary embolism by on-call radiology fellows and subsequently by radiology faculty. AJR 2004; 182:61 -66.
28. Verweij JW, Hofstee HM, Golding RP et al. Interobserver agreement between on-call radiology residents and radiology specialists in the diagnosis of pulmonary embolism using computed tomography pulmonary angiography. J Comput Assist Tomogr 2009; 33:952 -955.

Chapter 5

Wittenberg R
Peters JF
Weber M
Lely RJ
Cobben LPJ
Prokop M
Schaefer-Prokop CM

British J Radiol 2011 Dec (Epub ahead of print)

Standalone performance of a computer-assisted detection prototype for detection of acute pulmonary embolism: a multi-institutional comparison

Abstract

Purpose

To assess whether the performance of a computer-assisted detection (CAD) algorithm for acute pulmonary embolism (PE) differs in pulmonary CT angiographies (CTPA) acquired at various institutions.

Methods

In this retrospective study, we included 40 consecutive scans with and 40 without PE from 3 institutions (n=240) using 64-slice scanners of different manufacturers (General Electric; Philips; Siemens). CAD markers were classified as true or false positive (TP/FP) using independent evaluation by two readers and consultation of a third chest radiologist in discordant cases. Image quality parameters were subjectively scored using 4/5-point scales. Image noise and vascular enhancement were measured. Statistical analysis was done to correlate image quality of the 3 institutions with CAD standalone performance.

Results

Patient groups were comparable with respect to age (p=0.22), accompanying lung disease (p=0.12) and inpatient/outpatient ratio (p=0.67). On patient-basis, the sensitivity was 100% (34/34), 97% (37/38) and 92% (33/36) and the specificity was 18% (8/44), 15% (6/41) and 13% (5/39). Both did not significantly differ between the institutions (p=0.21 and p=0.820, respectively). The mean number of FP findings (4.5, 6.2 and 3.7) significantly varied (p=0.02 and p=0.03), but median numbers (2, 3 and 3) were comparable. Image quality parameters were significantly associated with the number of FP findings ($p<0.05$) but not with sensitivity. After correcting for noise and vascular enhancement the number of FP did not significantly differ between the 3 institutions (p=0.43).

Conclusions

CAD standalone performance is independent of scanner type but strongly related to image quality and thus scanning protocols.

Introduction

Multidetector computed tomography (MDCT) has become the first-line diagnostic imaging modality for pulmonary embolism (PE) at most institutions [1-4]. With the technical evolution of MDCT, an increasing number of ever thinner sections are generated that have to be systematically scrutinized for the presence of emboli in pulmonary arteries. This is a challenging task, especially for smaller and more peripheral pulmonary vessels. As a result, detection of subsegmental small emboli has a variable interobserver agreement. One study using a 16-slice MDCT scanner reported kappa values ranging from 0.56 to 0.85 [5].

The aims of automated computer-assisted detection (CAD) software for PE are to decrease perception errors, to reduce the workload and speed up evaluation, and to make reader performance less dependent of their level of skill or training. The majority of the previously published studies tested the standalone performance of various CAD algorithms [6-13]. Up to now there are four studies that test the influence of CAD on radiologists' detection performance [8, 14-16]. However, all studies so far tested the CAD on CT pulmonary angiograms (CTPA) from a single institution.

It is likely that the quality of the CTPA influences the performance of the CAD algorithm. Poorly timed contrast, motion artefacts, noise and the presence of accompanying lung diseases lead to higher numbers of FP findings [6, 7]. Because image quality depends on CT protocol parameters, the CT manufacturer, and patient instructions, it can be expected that the performance of a CAD algorithm may differ between institutions.

The purpose of this study was therefore to assess whether the performance of a CAD algorithm for acute PE differs in pulmonary CT angiographies acquired at various institutions.

Materials and Methods

Patient selection

In this institutional review board-approved study we retrospectively included 240 consecutive 64-slice CTPA scans from 3 institutions that use different manufacturers (Brilliance-64, Philips Medical Systems, Cleveland, OH, USA; GE LightSpeed Volume CT 64, Waukesha, Wisconsin, USA; Somatom Sensation 64, Siemens Medical Solutions, Forchheim, Germany). The first 40 consecutive scans with PE and 40 consecutive scans without PE were selected per institution between January and October 2008. Presence or absence of PE was decided upon their original radiology reports. All CTPA studies were obtained with the

Table 1. Computed Tomography protocols.

	Site A	Site B	Site C
Scan direction	cranio-caudal	caudo-cranial	cranio-caudal
kVp	120	120	100
Slice collimation (mm)	64 x 0.625	64 x 0.625	24 x 1.2
Rotation time (s)	0.4	0.4	0.37
Pitch	0.984	1.172	0.75
mAs	136	100	135
Contrast volume (ml)	70	90	40-70A
Concentration (iodine/ml)	320	300	300
Flow (ml/s)	5	5	6
NaCl chaser (ml)	20	30	30
Delay (s) / threshold to trigger (HU)	3	8 / 150	5-12
Reconstructed slice thickness (mm)	0.625	0.9	1.5
Reconstruction-interval (mm)	0.5	0.45	1.0

use of manufacturer-specific dose modulation software and local CT protocols (Table 1).

In total, 8 patients had to be excluded: in four patients the lung segmentation was mislead by air in abnormal locations leading to failure of the CAD algorithm, in 2 patients there were massive streak artefacts due to improper arm positions, 1 patient had chronic PE and 1 patient had a thrombus solely in the main artery, that is not included in the CAD analysis. This resulted in inclusion of 38, 39 and 38 positive CTPA and 40, 40 and 37 negative CTPA per site, respectively, according to the original reports.

Data analysis

All examinations were analysed by prototype CAD software (Philips Healthcare, Best, The Netherlands). After fully automatic segmentation of the lung and vessels [17, 18], the algorithm looks for contrast differences in both pulmonary arteries and veins >2mm and marks the difference if it exceeds a threshold of 150 HU. The main pulmonary arteries are not included in the CAD analysis. Candidate lesions detected by the CAD are indicated by a ROI and presented to the observer on demand. Processing time takes about 30 s per examination. The reference standard was established by consensus of at least two of three readings. All CT scans were independently evaluated by a researcher specially trained in reading PE studies (> 300 exams, RW) and a chest radiologist (> 15

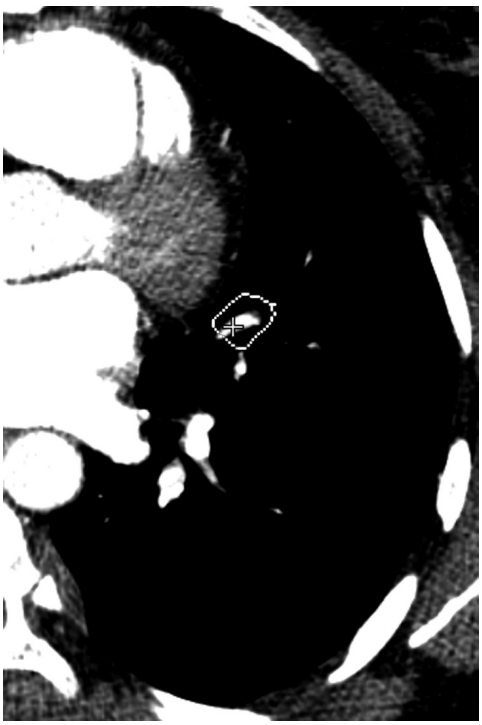

Figure 1. False positive finding found by CAD located in a vein.

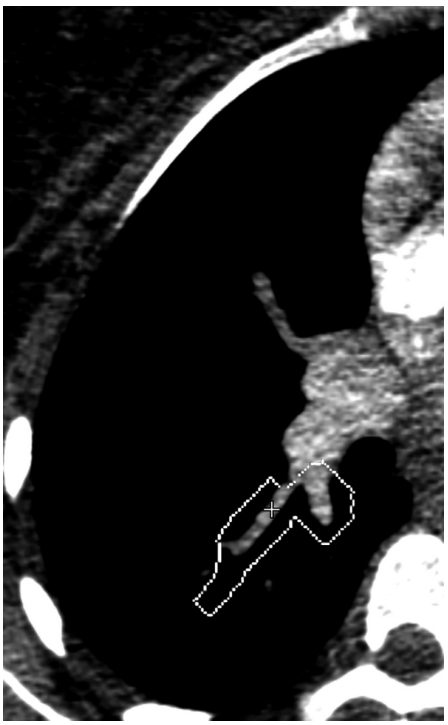

Figure 2. False positive finding found by CAD due to low vascular enhancement.

year experience, CMS). The readers first evaluated all datasets without CAD and subsequently the CAD results. They reported presence and anatomic locations of thrombi independently from each other without knowledge of the original reports. In case of discordant findings between these two readers, and/ or with the original report, a third experienced chest radiologist (> 15 years experience, MP) was consulted as an arbiter. The anatomic level of the thrombus was annotated as central, lobar, segmental or subsegmental according to its proximal end. The CAD findings were compared with this reference standard and classified as true positive (TP) or false positive (FP). We differentiated central locations in main and lobar arteries from peripheral locations in segmental and subsegmental arteries. Underlying reasons for FP CAD markers were classified as related to anatomic structures (veins, lymphatic tissue and intrapulmonary opacities) or related to motion artefacts and low vascular enhancement (Figures 1 and 2).

Image noise was measured in the descending aorta at the level of the bifurcation of the trachea as standard deviation (SD) of CT numbers using a standardized ROI of 1 cm^2. Noise per scan was calculated as average over 3

measurements. To evaluate the vascular enhancement we measured the mean CT number in Hounsfield Units using individual ROIs at the level of the central pulmonary artery (ROI with 2 cm diameter), the segmental (3 mm) and the subsegmental arteries (1 mm) in the left upper and lower lobe. The enhancement in the subsegmental artery in the upper lobe was measured at the level of the aortic arch and in the lower lobe between the inferior pulmonary veins and just above the diaphragm. In each artery we measured the enhancement 3 times and calculated the average. If for any reason (e.g. embolus, underlying lung disease, motion artefacts) the enhancement in the left lung could not be measured, equivalent arteries in the right lung were selected.

The quality of the CT examinations were subjectively scored by a researcher (RW) using a 4- and 5-point scale, respectively, with respect to the overall quality, the presence of motion artefacts and the effects of accompanying lung disease (Table 2).

Table 2. Subjective scores of image quality characteristics.

Scale and score	Description
Overall image quality	
1	Inadequate, no diagnosis of PE[A] possible
2	Low, diagnosis possible until the lobar level
3	Sufficient, diagnosis possible until the segmental level
4	Good, diagnosis possible until the subsegmental level
5	Excellent
Motion artefacts	
1	Massive, no diagnosis possible
2	Definite, establishment of diagnosis impeded
3	Moderate, but image sufficient for diagnosis of PE
4	Minor
5	None
Accompanying lung disease	
1	Present, lung disease disturbing for CAD[B] and radiologist
2	Present, lung disease disturbing for CAD
3	Present, but no influence
4	None

[A] Pulmonary embolism.
[B] Computer-assisted detection

Statistical analysis

Statistical analysis was performed with SPSS (version 15.0). For all tests a p-value < 0.05 was considered significant. Differences in patient groups were assessed using a chi square test with respect to sex and in-/outpatient ratio and an ANOVA test with respect to patient age.

Sensitivity and specificity for the presence of PE were calculated on per-patient basis separately for each institution. Furthermore, the sensitivity was calculated on per-lesion basis. A Fisher Halton Freeman test was used to assess differences between the 3 institutions.

We used factor analysis and Cronbach's Alpha statistics to assess the variability between the vascular enhancement measured at the central, lobar, segmental and subsegmental level. An ANOVA followed by a Hochbergs GT2 post hoc test were used to test for significance of difference with respect to the vascular enhancement and noise between the 3 institutions.

A Kruskal Wallis test followed by a post hoc Shaffer corrected Mann-Whitney U-test was used to test for significance of difference between the 3 institutions with respect to overall quality, motion artefacts, presence of accompanying lung disease and number of FP findings. An ANCOVA with noise and vascular enhancement as covariates was performed to establish the relation between scanner type and number of FP findings.

To assess the correlation between sensitivity per scan and noise, vascular enhancement and overall image quality a Pearson correlation or Spearman's rank correlation test was used, respectively. To assess the correlation between the various image quality parameters and the number of FP a multiple linear regression analysis was applied.

Results

Study groups

The three patient groups did not significantly differ with respect to age (p = 0.220) and inpatient/outpatient ratio (p = 0.674) (Table 3). The reference standard differed from the original reports in 9 patients: 1 patient originally reported as negative, the standard determined as positive and 8 patients originally reported as positive, the standard determined as negative. Thus, the reference standard for the three institutions rated 34, 38, and 36 scans positive for PE and 44, 41 and 39 scans negative for PE, respectively.

There were on average 6 (range 1-18) thrombi per patient in institution A, 5 (range 1-18) thrombi per patient in institution B and 4 (range 1-17) thrombi per patient in institution C.

Table 3. Results: Patient group characteristics.

	Site A	Site B	Site C
Mean age (range in years)	58 (18-88)	61 (27-93)	63 (11-91)
Inpatients (%)	51	58	56
Sex f:m (%)	68:32	48:52	47:53
Number of positive scans	34	38	36
Number of negative scans	44	41	39

Table 4. Results: Computer-assisted detection performance for the 3 different sites.

	Site A	Site B	Site C
Sensitivity on a per-patient basis (%)	100	97	92
Sensitivity on a per-lesion basis (%)	76	75	64
Specificity	18	15	13

Sensitivity and specificity

The sensitivity on per-patient basis was not significantly different between the three institutions with 100% (34/34), 97% (37/38) and 92% (33/36), respectively (p=0.21).

The sensitivity on per-lesion basis was significantly different with 76% (165/216), 75% (146/194) and 64% (84/132), respectively (p=0.025). CAD found in total 16 out of the 17 patients (94%) with only isolated subsegmental emboli.

The specificity of CAD on per-patient basis was not significantly different between the three institutions with 18% (8/44), 15% (6/41) and 13% (5/39), respectively (p=0.820) (Table 4).

Analysis of false positives

The mean number of false positive CAD findings per patient was 4.5 (median 2, range 0-29), 3.7 (median 3, range 0-20) and 6.2 (median 3, range 0-23), respectively, with the latter being significantly different from the other two (p=0.021 and p=0.03). In most scans (63-75%) CAD found 5 or less FP candidates. After correcting for differences of noise and vascular enhancement using an ANCOVA analysis, the mean number of FP findings per patient did not significantly differ between the three institutions (p=0.425). In all institutions most of the FP findings were located in veins or intrapulmonary opacities (Table 5).

Table 5. Results: Analysis of false positive computer-assisted detection findings.

	Site A	Site B	Site C
Mean number of FP	4.5	6.2	3.7
Median number of FP	2	3	3
≤ 5 CAD findings per scan (%)	72	63	75
> 10 CAD findings per scan (%)	17	22	8
Veins as FP finding (% of total FP)	27	34	37
Intrapulmonary opacifications as FP finding (% of total FP)	26	17	29

Vascular enhancement and noise

A factor analysis, performed for each hospital separately, revealed a high correlation between the mean vascular enhancement of the central, segmental and subsegmental arteries allowing for calculating a single average enhancement measure per institution. This amounted to 384 HU, 266 HU and 429 HU, respectively (Table 6, Figure 2) with all differences being significant at pair wise comparisons ($p<0.001$ to $p=0.039$). There was a significant negative correlation between the number of FP and vascular enhancement in two sites ($p=0.023$ and $p=0.046$) but no significant correlation between vascular enhancement and the sensitivity per scan ($p=0.872$).

The mean noise amounted to 34.4 HU, 30.2 HU and 34 HU, respectively (Table 6) with the differences between the lowest noise level and the remaining two being significant at $p=0.001$ and $p=0.004$. There was a significant positive correlation between the number of FP and the noise in two sites ($p=0.029$ and $p=0.012$) but no significant difference between the noise and the sensitivity per scan ($p=0.726$).

Overall quality, motion artefacts and accompanying lung disease

There were no significant differences between the 3 sites with respect to the presence and effects of accompanying lung diseases ($p=0.123$) and the severity of motion artefacts ($p=0.356$). There was a significant positive correlation between the number of FP and the severity of motion artefacts in two sites (multiple regression analysis, $p=0.047$ and $p=0.002$).

The overall quality was significantly lower at site B than at site A ($p=0.011$) and C ($p=0.002$) (Table 6). There was a significant correlation between the number of FP and the overall quality in two sites ($p<0.001$ and ($p=0.001$) but no significant difference between the overall quality and the sensitivity per scan ($p=0.647$).

Table 6. Results: Quality parameters for the 3 different sites.

	Site A	Site B	Site C
Vascular Enhancement (HU)	384	266	429
Noise (HU)	34.4	30.2	34
Overall quality (scale 1-5[A])	4.3	3.9	4.4
Accompanying lung disease (scale 1-4[A])	3	2.8	2.8
Motion artefacts (scale 1-4[A])	3.1	3	2.9

[A] The higher the score the better the image quality and the lower the presence and influence of overall quality, motion artefacts and accompanying lung disease, respectively.

Discussion

There is little known about the performance of CAD software when used in different institutions and whether results made in one site can be transferred to another site. In addition to patient-related factors such as concomitant lung disease or motion artefacts, image quality is influenced by the CT data acquisition protocol that may vary per institution. We therefore investigated the standalone performance of a CAD prototype in 3 institutions with respect to sensitivity, specificity and a number of subjectively and objectively scored image parameters. Both performance measures, sensitivity and specificity, were assessed on patient basis and lesion basis. We consider the first the more important measure from a clinical point of view, given the fact that detection of an isolated embolus is sufficient to call a study positive for the presence of PE and thus to indicate need for treatment. The latter, however, we consider more suited to describe the level of standalone performance of this prototype software.

For CAD software to be beneficial in clinical practice as a second reader, a high sensitivity appears warranted. In our study we calculated both, the sensitivity on a per-lesion basis as a criteria of the software's performance and on a per-patient basis, which appears to be clinically more relevant given the current concept that in hemodynamically stable patients treatment is based on a yes/no decision for the presence of PE .

In our study, the sensitivity on a per-patient basis varied between 92% and 100%, thus yielded a performance in all three institutions that compares favourably with published data so far reporting a sensitivity of 53%, 86% and 94% [8, 10, 11]. Statistically, the difference did not reach significance which does not necessarily mean that the observed difference is not noteworthy. The lower sensitivity in one of the three sites was due to false negative findings in 3 of the 36 patients found positive for PE. Two patients had emboli in only one location

and 1 patient in 3 locations. The missed intravascular defects were of various sizes and on various anatomic levels (1 lobar, 1 segmental and 3 subsegmental). Underlying reasons for the failures included low vascular enhancement in 2 and perivascular pulmonary consolidations in one patient.

On a per-lesion basis the sensitivity differed significantly between the three institutions. However, as opposed to a significant correlation between the number of FP findings and parameters of image quality, we could not find a significant correlation between the sensitivity and parameters of image quality. This was further underlined by the finding that the institution with the lowest sensitivity did not show a lower mean vascular enhancement or an increased noise. We concluded from these findings that the relationship between sensitivity and image quality seems to be more complex involving also factors relating to the locations and multitude of the emboli. This finding is conform with previously reported results by Dewailly et al [13] that also could not find an effect of overall image quality or different scanning conditions on the detection rate of peripheral clots of a prototype CAD that was different from the one we tested.

In the entire study group, CAD found 16 out of 17 (94%) patients with isolated subsegmental emboli. This is noteworthy, because usually radiologists securely detect the rather obvious central or lobar emboli, but may miss isolated subsegmental emboli. It has to be noted, however, that from a clinical point of view the significance of isolated subsegmental emboli is still uncertain. In the literature it is suggested that the clinical relevance of subsegmental emboli is larger in individuals with cardiopulmonary restriction and that subsegmental emboli may predict a more severe emboli disease in the future or the development of pulmonary hypertension [19].

On a per-patient basis, we found generally low specificities of 18%, 15% and 13%, respectively. For 2 out of the 3 institutions each we found a significant correlation between the number of FP findings and the overall quality, vascular enhancement, noise and motion artefacts, respectively. For the remaining third institution, the numbers of examinations with low vascular enhancement or high image noise, respectively, were too small to proof a statistically significant correlation with the number of FP findings. This, however, does not mean that a correlation exists also in this institution.

On average, CAD showed 4.5, 6.2 and 3.7 FP findings per patient, respectively, with the difference being significant for the site with the highest mean number of FP. Though the mean numbers of FP findings differed, the median numbers of FP were comparable and amounted to 2, 3 and 3, respectively. This suggests that a subgroup of examinations with exceedingly high numbers of FP candidates were responsible for the significant increase of the mean FP findings and not

a generally lower image quality of all examinations. If we would consider a maximum of 5 FP candidates per scan an acceptable threshold for clinical application, this criterion was fulfilled in 72%, 63% and 75% of examinations, respectively. In all three institutions most false positive candidates were localised in veins. It has to be noted that all false positive candidates located in veins were classified as such based on the misinterpretation of the anatomic location and irrespective of other aspects such as inhomogeneous vascular enhancement, a contrast difference against a surrounding pulmonary consolidation or a motion artefact contributed to the misinterpretation of the CAD algorithm. It is therefore likely that multiple conditions have contributed to the high number of false positive CAD candidates in veins. Simply adjusting the protocol to achieve high venous contrast would only partially solve the problem. Teaching the algorithm to differentiate between veins and arteries seem to be a more promising approach. This is an important aspect considering that a too high number of FP findings will result in prolongation of reading time and may decrease confidence in true positive findings of the CAD. Therefore further lowering the number of false positives, especially in examinations of lower image quality, seems warranted before applying CAD into clinical routine.

It is important to note that after ANCOVA analysis, correcting for noise and vascular enhancement, we did not find a significant difference between the institutions with respect to the number of FP candidates. We concluded from this that only image quality parameters and thus CT protocols are responsible for the differences in number of FP CAD candidates between the institutions and not the scanner type used.

Our study has the following limitations. Per institution, we consecutively included the first 40 scans reported positive for PE and the first 40 scans reported negative for PE. As intended, that way scans were included irrespective of image quality, time of acquisition or patient condition. We chose for almost equal numbers of positive and negative scans to get a realistic estimation of the performance of the CAD software in both groups of examinations. The standalone performance of the CAD was determined by a reference that had been established by independent readings of two and in cases of discordance by assessment of a third experienced radiologist. Though this reference has to be considered as not optimal it reflects clinical applicability.

In summary, the CAD prototype yields a comparable standalone performance with different scanner types and various CTPA protocols if image quality matches. We found a significant correlation between image quality parameters and the number of FP suggesting that an optimisation of the protocol with respect to vascular enhancement, noise and the avoidance of breathing artefacts is useful to increase the specificity of the CAD. After correcting for image

quality parameters, the scanner type had no significant influence on the number of FP candidates. Though high on patient basis in all three institutions with numbers between 92 and 100%, significant differences in sensitivity existed between the three institutions on per-lesion basis. These differences appeared not only be influenced by image quality but also by a combination of patient related aspects, for which further analysis is needed to determine magnitude and clinical relevance.

Reference List

1. British Thoracic Society guidelines for the management of suspected acute pulmonary embolism. Thorax 2003; 58(6):470-83.
2. Cronin P, Weg JG, Kazerooni EA. The role of multidetector computed tomography angiography for the diagnosis of pulmonary embolism. Semin Nucl Med 2008; 38(6):418-31.
3. Remy-Jardin M, Pistolesi M, Goodman LR et al. Management of suspected acute pulmonary embolism in the era of CT angiography: a statement from the Fleischner Society. Radiology 2007; 245(2):315-29.
4. Schaefer-Prokop C, Prokop M. MDCT for the diagnosis of acute pulmonary embolism. Eur Radiol 2005; 15 Suppl 4:D37-D41.
5. Brunot S, Corneloup O, Latrabe V et al. Reproducibility of multi-detector spiral computed tomography in detection of sub-segmental acute pulmonary embolism. Eur Radiol 2005; 15(10):2057-63.
6. 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(6):651-8.
7. Schoepf UJ, Schneider AC, Das M et al. Pulmonary embolism: computer-aided detection at multidetector row spiral computed tomography. J Thorac Imaging 2007; 22(4):319-23.
8. Walsham AC, Roberts HC, Kashani HM et al. The use of computer-aided detection for the assessment of pulmonary arterial filling defects at computed tomographic angiography. J Comput Assist Tomogr 2008; 32(6):913-8.
9. Zhou C, Chan HP, Patel S et al. Preliminary investigation of computer-aided detection of pulmonary embolism in three-dimensional computed tomography pulmonary angiography images. Acad Radiol 2005; 12(6):782-92.
10. Maizlin ZV, Vos PM, Godoy MC et al. Computer-aided detection of pulmonary embolism on CT angiography: initial experience. J Thorac Imaging 2007; 22(4):324-9.
11. 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 2009; 20(4):801-6.
12. Masutani Y, MacMahon H, Doi K. Computerized detection of pulmonary embolism in spiral CT angiography based on volumetric image analysis. IEEE Trans Med Imaging 2002; 21(12):1517-23.
13. Dewailly M, Remy-Jardin M, Duhamel A et al. Computer-aided detection of acute pulmonary embolism with 64-slice multi-detector row computed tomography: impact of the scanning conditions and overall image quality in the detection of peripheral clots. J Comput Assist Tomogr 2010; 34(1):23-30.
14. Das M, Mühlenbruch G, Helm A et al. Computer-aided detection of pulmonary embolism: Influence on radiologists‘ detection performance with respect to vessel segments. Eur Radiol 2008;18: 1350-5.

15. Engelke C, Schmidt S, Bakai A et al. Computer-assisted detection of pulmonary embolism: performance evaluation in consensus with experienced and inexperienced chest radiologists. Eur Radiol 2008; 18(2):298-307.
16. Engelke C, Schmidt S, Auer F et al. Does computer-assisted detection of pulmonary emboli enhance severity assessment and risk stratification in acute pulmonary embolism? Clin Radiol 2010; 65(2):137-44.
17. Buelow T, Wiemker R, Blaffert T et al. Automatic extraction of the pulmonary artery tree from multi-slice CT data. Spie 2005; 5746:730-40.
18. Bouma H, Sonnemans JJ, Vilanova A et al. Automatic detection of pulmonary embolism in CTA images. IEEE Trans Med Imaging 2009; 28(8):1223-30.
19. Schoepf UJ, Costello P. CT angiography for diagnosis of pulmonary embolism: state of the art. Radiology 2004; 230(2):329-37.

Part 3

Impact of a computer-assisted detection prototype used as a second or a concurrent reader on readers with varying experience

Chapter 6

Wittenberg R
Berger FH
Peters JF
Weber M
van Hoorn F
Beenen LFM
van Doorn MMAC
van Schuppen J
Zijlstra IJAJ
Prokop M
Schaefer-Prokop CM

Radiology 2012 Jan: 262(1):305-13

Acute pulmonary embolism: effect of a computer-assisted detection prototype on diagnosis - An observer study

Abstract

Purpose

The purpose of this study was to assess the impact of a computer-assisted detection (CAD) prototype on observer performance for detection of acute pulmonary embolism (PE) in computed tomography pulmonary angiography (CTPA).

Methods and Materials

In this institutional review board-approved retrospective study, six observers of varying experience evaluated 158 negative and 51 positive CTPA (mean age 57 years; 111 women), consecutively obtained during off-hours. Observers were asked to determine PE and to rank their diagnostic confidence without and subsequently with CAD as second reader within a single reading session. Reading time was separately measured for both readings. Reader data were compared with an independent standard established by two readers and a third chest radiologist in case of discordant results. Statistical evaluation was performed on patient basis using logistic regression for repeated measurements and Pearson correlation.

Results

With CAD, there was a significant increase of readers' sensitivity ($p=0.014$) without loss of specificity ($p=0.853$) on a per patient basis. CAD assisted the readers to correctly determine PE in 15 cases with the most proximal embolus at the segmental level in 4 cases and at the subsegmental level in 11 cases. In 8 cases, readers accepted false positive (FP) CAD candidates on scans negative for PE and in one case, a reader dismissed a true positive finding. Reading time was on average extended by 22s using CAD.

Conclusions

At the expense of increased reading time, CAD has the potential to increase reader sensitivity for detecting segmental and subsegmental PE without significant loss of specificity.

Introduction

In the past decade, computer-assisted detection (CAD) software has been investigated for various radiologic applications, including mammography, thoracic computed tomography (CT) and virtual CT colonoscopy [1-3]. Contrast enhanced multidetector computed tomography (MDCT) is the currently accepted method of choice for the detection of acute pulmonary embolism (PE) [4-6]. However, analysing the complex pulmonary vasculature to the subsegmental level to detect PE in the large data sets generated with multidetector row CT examinations is a demanding task. CAD algorithms for the detection of intravascular filling defects were developed to aid radiologists with detection of PE and are being investigated with respect to their effect on diagnostic accuracy, performance homogeneity and time efficiency.

Most studies in the literature have only evaluated the standalone performance of CAD algorithms [7-15]. Whether CAD can improve radiologists' performance in clinical practice will depend not only on the accuracy of the system, but also on other factors, such as radiologists' experience, their confidence in the software and their ability to differentiate true- from false positive candidate lesions. To our knowledge, researchers in four studies have evaluated the effect of CAD on reader performance for the detection of PE [13, 16-18]. In two [17,18] of these four studies, investigators found an increase in readers' sensitivity when they used CAD, but both studies were based on rather small patient groups (n=56 and 58) and included only studies positive for PE. In clinical practice, however, since the majority of scans are obtained to exclude PE, it is also important to show the effect of CAD on readers' performance with respect to negative scans in larger patient groups.

The purpose of our study was to assess the effect of a CAD prototype on observer performance for detection of acute PE on CT pulmonary angiographic images.

Materials and Methods

CAD software

The study was supported by a research grant from Philips Healthcare (Best, the Netherlands). Neither Philips Healthcare nor the author who is an employee of Philips Healthcare (J.F.P.) had control of inclusion of data or an influence on analysis of data.

In this study, we tested a prototype CAD algorithm (Philips Healthcare). The reading took place at a workstation with special viewing software (Easyscil;

Philips Healthcare). We used 1280x1024–pixel color liquid crystal display monitors (Brilliance 190P; Philips Healthcare) that were equipped with a calibration tool to assure Digital Imaging and Communications in Medicine conformity. Ambient lighting in the reading room was dimmed. The readers were allowed to adjust the window and level settings to their references and could zoom and pan the images. Technical information for the CAD algorithm is given elsewhere in more detail [14]. In summary, detection of candidate lesions with CAD is based on a gray-value analysis of the cross-sectional images perpendicular to the centerlines of the blood vessels [19,20]. Features for identifying PE include completely occluded stretches of blood vessels as well as contrast-to-tissue transitions. The CAD analysis does not differentiate between veins and arteries and does not find emboli in the main pulmonary arteries. Candidate lesions identified with CAD as potential PE are clustered and presented to the radiologist as colored markers added to the original data set on demand.

Study design

In this institutional review board–approved study, we retrospectively included 215 consecutive CT pulmonary angiographic studies performed in a university hospital during night shifts and on weekends between May 2007 and May 2008. Identifying patient information was removed from the studies. All patients had been referred to the radiology department for CT pulmonary angiography because they were suspected of having acute PE. An automatic injector (Stellant Dual CT Injector; Medrad Europe, Beek, the Netherlands) was used for intravenous bolus injection of iopromide (Ultravist 300; Schering, Berlin, Germany) (Table 1). All scans used in this study were taken from a set of consecutive scans that had been previously evaluated to determine the standalone performance of CAD and to assess the effect of image quality on CAD performance [14,15].

Six patients were excluded from further evaluation because the CAD algorithm failed to analyse these data sets owing to segmentation errors. In one case, the CAD algorithm failed because the patient had a pneumothorax. The other five patients had a trachea tube for respiratory ventilation, which resulted in a connection between extrathoracic and intrapulmonary air and subsequent failure of the segmentation of trachea and lungs.

The final study group consisted of 209 patients, with a mean age of 57 years (median, 59 years; range, 18–87 years). There were 98 men (mean age, 58 years; median, 60 years; range, 18–87 years) and 111 women (mean age, 56 years; median, 58 years; range, 18–87 years). One hundred nineteen patients were imaged with a 16-detector row scanner (MX 8000 IXDT or Brilliance-16; Philips Medical Systems, Cleveland, Ohio), and 90 patients were imaged with a 64-detector row scanner (Brilliance-64; Philips Medical Systems). There were

Table 1. CT protocols used on the 16- and 64-detector-row scanner

	16-detector-row	64-detector-row
Scan direction	caudo-cranial	caudo-cranial
kVp	90	120
Collimation (mm)	0.75	0.065
Rotation time (s)	0.5	0.4
Pitch	1.188	1.172
mAs	180	100
Volume (ml) of contrast	90	90
Concentration (mg iodine/ml)	300	300
Flow of injection (ml/s)	4.0	5.0
Saline chaser (ml)	40	30
Threshold of scan triggering (HU)	150	150
Reconstructed slice thickness (mm)	1.0	0.9
Reconstruction-interval (mm)	0.5	0.45

104 inpatients (including 10 patients from the intensive care unit), 102 patients from the emergency department and three patients from the outpatient clinic.

A reference standard with respect to the localisation of intraarterial intravascular defects was established by the independent evaluations of a researcher specially trained in reading for PE (RW, with experience of performing more than 400 examinations) and a chest radiologist (CMS, with more than 15 years experience). In cases of discordant findings between the two readers and/or with the original report, a third chest radiologist (MP, with more than 15 years of experience) was consulted as an arbiter. This expert panel noted the proximal margin of the most proximal embolus per patient. Standard nomenclature, derived from Boyden [21] and from Jackson and Huber [22], was used to identify the main, lobar, segmental and subsegmental structures. The proximal margin of the most proximal embolus was located at the main or lobar level in 28 cases and at the segmental level in 11 cases. In 12 cases, only subsegmental emboli were found. Subsequently, the readers who established the reference standard reviewed the results of CAD and noted the number of false positive findings. According to the reference standard, 24% (51 of 209) of the studies were positive for PE, and 76% (158 of 209) were negative for PE.

Reading methodology

Six readers with different levels of experience (reader 1, fellowship-trained radiologist; reader 2, 4th-year resident; reader 3, 3rd-year resident; reader 4, 2nd-

year resident; reader 5, emergency radiologist; and reader 6, 2nd-year resident) independently read all 209 scans while blinded to the clinical reports. Within one reading session, the readers assessed the study scans for the presence of PE, first without results of the CAD analysis and immediately afterward with disclosure of the CAD candidate lesions. To approximate the clinical situation as closely as possible, how much time to devote to the CAD results was left up to the readers. Eight training cases, which were not part of the study, were used to familiarise the readers with the software. After the initial unassisted evaluation, the CAD candidate lesions were made visible for second reading. Readers recorded the time for the initial reading and the extra time for the subsequent review of the CAD candidate lesions. The CAD candidate lesions were invisibly preprocessed during the initial reading; therefore, processing time was not an issue and was not measured. All readers were asked to separately record their diagnosis without CAD and that with CAD with respect to the diagnosis of PE by using a nine-point scale to indicate their level of confidence (score 1, PE definitely not present; score 9, PE definitely present). Readers were not asked to specifically document all emboli they localised but only to determine the presence of PE. Finally, they were asked to subjectively score the image quality of the scans by using a five-point scale (Table 2).

Statistical analysis

All statistics were calculated on a per-patient basis. Sensitivities and specificities without CAD and those with CAD for all readers were compared by using logistic regression for repeated measurements, which corrects for multiple readers evaluating the same cases. The terms used in our model were method (without CAD and with CAD), reader and reader-method interaction. Sensitivity without CAD and that with CAD were also compared on an individual reader basis by using the McNemar test. All analyses were evaluated on the basis of the readers final diagnosis as compared with the reference standard. A receiver operating

Table 2. Definition of subjective image quality scores.

Scale and score	Description
Overall image quality	
1	Inadequate, no diagnosis of PE possible
2	Low, diagnosis possible to the lobar level
3	Sufficient, diagnosis possible to the segmental level
4	Good, diagnosis possible to the subsegmental level
5	Excellent

characteristic curve analysis was used for comparison of observer performance in the detection of PE with CAD and that without CAD. The areas under the receiver operating characteristic curve were compared by using software (MedCalc for Windows, version 11.2.1.0; Med Calc Software, Mariakerke, Belgium) [23].

Since confidence and quality scores are assumed to be equidistant, they were treated as metric data. A one-way repeated analysis of variance followed by a Bonferroni post hoc test was used to test for differences between readers with respect to the quality scores and the extra time needed to use CAD. A Pearson correlation was used to test for correlations between the mean number of false positive CAD candidate lesions per scan and the confidence score and between the mean reading time and the quality score, both with CAD and without CAD. A two-way repeated analysis of variance was used to assess the effect of CAD on readers' confidence scores.

Results

Sensitivity, specificity and number of FP candidates

The standalone performance of the CAD algorithm on a per-patient basis amounted to a sensitivity of 96% (49/51) and a specificity of 22% (35/158).

Readers' sensitivity and specificity with and without CAD were calculated on a per-patient basis. Sensitivity improved with CAD for all 6 readers (reader

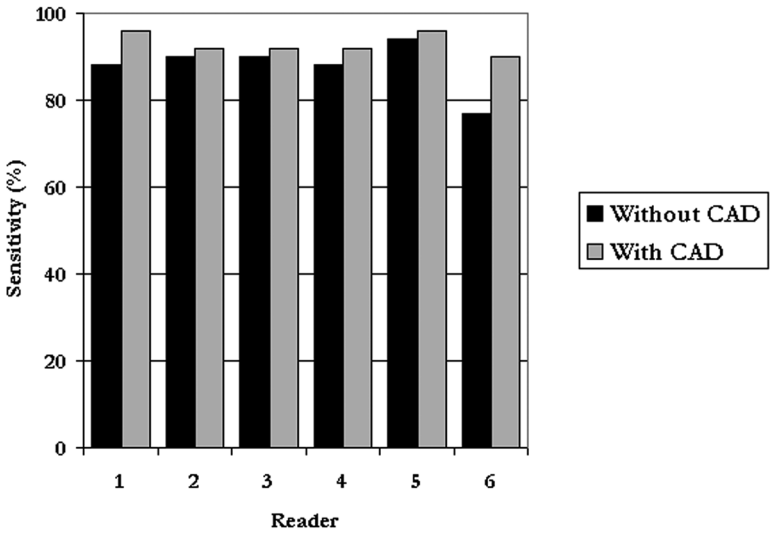

Figure 1. Sensitivity with and without using CAD for 6 readers.

1 88% (45/51) to 96% (49/51); reader 2 90% (46/51) to 92% (47/51); reader 3 90% (46/51) to 92% (47/51); reader 4 88% (45/51) to 92% (47/51); reader 5 94% (48/51) to 96% (49/51); and reader 6 77% (40/51) to 90% (46/51)) (Figure 1).

Logistic regression for repeated measurements, taking all reader data into account, revealed a significant increase in readers' sensitivity when using CAD (p=0.014 using the term CAD. The interaction between method and reader was not significant (p=0.140) implying that the gain in sensitivity was appreciated equally by all readers. The difference in sensitivities with and without CAD remained significant after excluding reader 6, who had shown the highest increase in sensitivity with CAD, from the data analysis (p=0.042). Analysis of individual reader data separately using the McNemar test showed a significant difference only for reader 6 (p=0.016).

Of a total of 306 (6x51) scans positive for PE CAD assisted the readers to beneficially change their diagnosis from false negative to true positive fifteen times (4.9%): 4 times the most proximal embolus was at the segmental level and 11 times at the subsegmental level (Figure 2). These 15 beneficial changes of diagnosis referred to 10 different patients with PE, with seven patients for whom only one reader changed the diagnosis after using CAD, one patient for whom two readers changed the diagnosis and two patients for whom three readers changed the diagnosis.

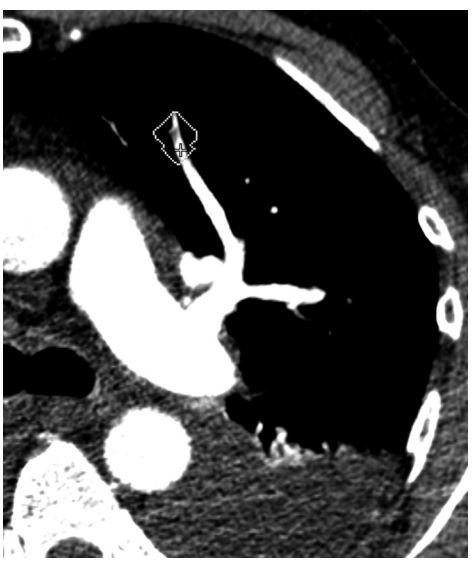

Figure 2. An example of an intra-arterial defect that had been primarily missed but assigned to the correct diagnosis with the assistance of CAD.

Specificity changed for readers 1 from 89% (141/158) to 91% (144/158)) and for reader 2 from (97% (153/158) to 96% (152/158). Specificity remained constant for readers 3 (96%, 152/158), 4 (93%, 147/158), 5 (98%, 155/158) and 6 (98%, 155/158) without and with CAD (Figure 3). There was no significant difference in readers' specificity without CAD or with CAD (p=0.853). The difference in the interaction between method and reader was also not significant (p=0.444), implying that the gain in specificity was appreciated equally by all readers.

Of a total of 948 (158 cases for six readers) viewed scans with negative reference standard findings for PE, readers detrimentally changed their diagnosis eight times (<1%) from a true negative to a false positive diagnosis with the availability of the CAD candidate lesions. These eight detrimental changes of diagnosis referred to seven different patients without PE, with six patients for whom only one reader changed the diagnosis after using CAD and one patient for whom two readers changed the diagnosis. In one case, a reader dismissed a true positive finding (Figure 4). There was no clear relationship between these results and the level of experience of the reader.

CAD indicated a mean of 4.9 false positive candidate lesions per scan (median, three; range, 0 to 42). Most of the false positive CAD candidate lesions were in veins (32%, 333/1030), followed by low opacified arteries (20%, 205/1030). In 17% (170/1030), false positive candidate lesions were caused by motion artefacts.

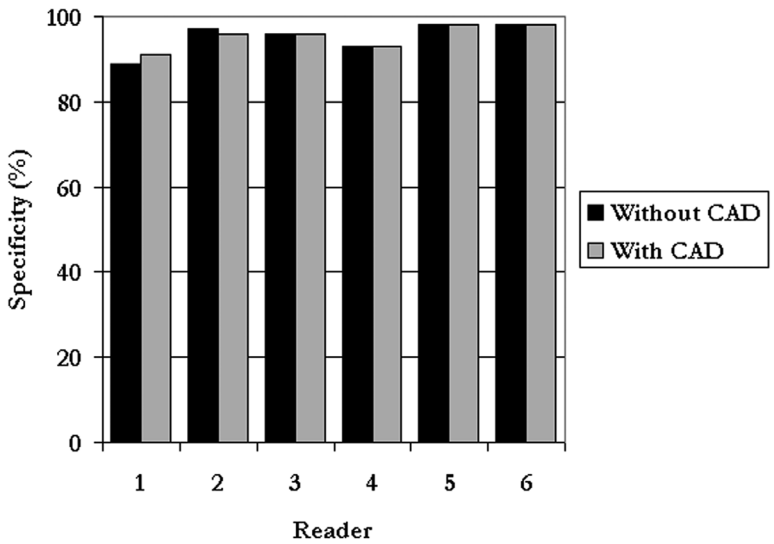

Figure 3. Specificity with and without using CAD for 6 readers.

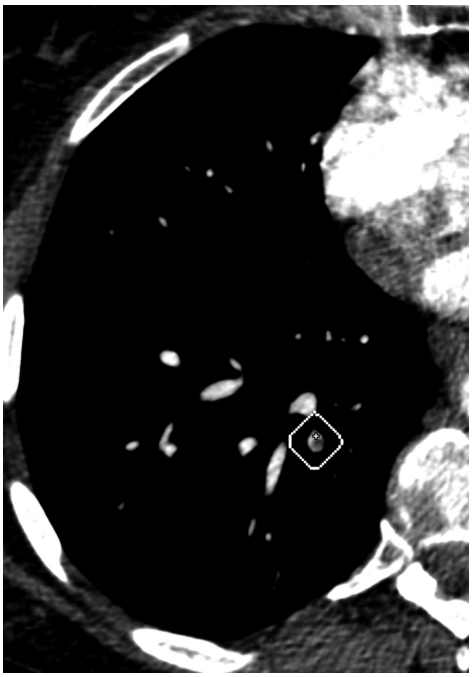

Figure 4. An example for a true positive CAD candidate but dismissed by one reader.

Receiver operating characteristic curves

A significant increase in the area (A_z) under the receiver operating characteristic (ROC) curve was not found for any of the readers. The areas under the ROC curve without CAD and with CAD, respectively, amounted to 0.908 and 0.951 for reader 1 (p=0.29); 0.97 and 0.99 for reader 2 (p=0.35); 0.96 and 0.97 for reader 3 (p=0.73); 0.96 and 0.97 for reader 4 (p=0.44); 0.99 and 0.99 for reader 5 (p=0.7); and 0.93 and 0.97 for reader 6 (p=0.18).

Confidence and quality scores

There was a significant increase in confidence scores when using CAD (p<0.001), though the effect of CAD was not the same for all readers (p<0 .001). Without CAD, the confidence score for all readers was significantly positively correlated with a shorter reading time (p<0.008, -0.184 ≤ r ≥ -0.459). For all readers, there was a significant positive correlation between the mean confidence and quality scores with CAD (p<0.001, r=0.842) and without CAD (p<0.001, r=0.839).

Extra reading time

The mean reading time per examination without CAD was 75 seconds per

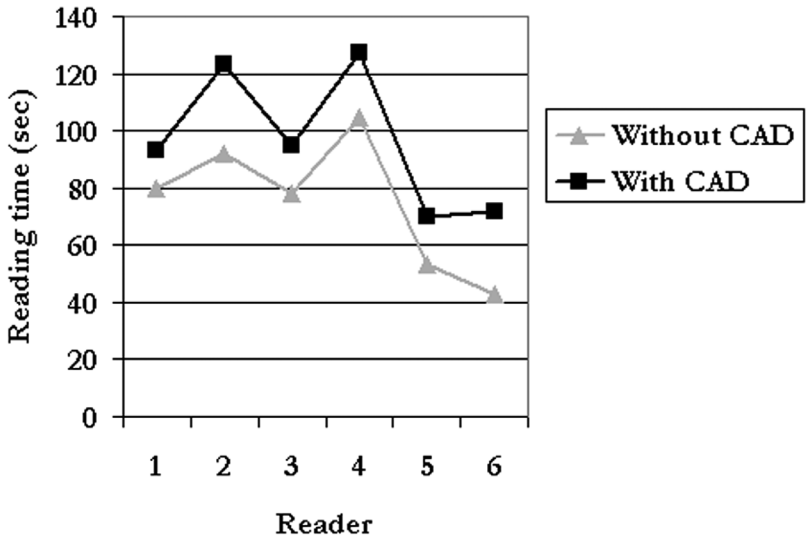

Figure 5. Mean reading time with and without CAD: There were significant differences between readers with respect to the length of unassisted reading times as well as to the length of extension of reading times with CAD

reader (range, 43-105 seconds per reader). The mean extra time needed to use CAD was 22 seconds per reader (range, 13-31 extra seconds per reader) (Figure 5). There were significant differences among readers with respect to the length of reading time ($p<0.001$). There was a significant positive correlation between the mean number of false positive CAD candidate lesions and the mean extra time ($p<0.001$, $r=0.402$). For the 15 cases in which readers took a diagnostic advantage, the mean extra time used was 62 seconds per scan (range, 18-158 seconds per scan).

Discussion

Lately, CAD algorithms have been developed to assist radiologists in decision making and to increase their performance for the detection of acute PE. To our knowledge, only four studies [13, 16-18] have been published that show the effect of CAD on readers for the detection of PE. In these studies, relatively small patient groups (n=43, 56, 58 and 118, respectively) or only selected examinations were included and researchers reported controversial results. In our study, we included a larger number of consecutive scans to evaluate the effect of a CAD prototype on the performance of both radiologists and residents. All patients were scanned during night or weekend shifts, thus representing a group of

examinations that were obtained with potentially suboptimal conditions and are typically reviewed by readers of varying experience.

Our results using logistic regression for repeated measurements showed a mean increase in sensitivity with CAD for the detection of segmental and subsegmental emboli across readers without a significant loss in specificity. The extent of the differences in sensitivity and specificity did not vary significantly between readers. Also, after excluding the data of reader 6, whose performance had improved the most with the use of CAD (sensitivity increase from 78% to 90%), statistical analysis still yielded a smaller yet significant difference. Data analysis of the readers separately by using the McNemar test revealed a significant difference only for reader 6. This seeming discordance is likely to be explained by the fact that the effect of CAD on reader performance is based on a small subset of examinations within the study group, namely those patients with peripheral (segmental and subsegmental) low conspicuous emboli. On an individual basis, differences were too small to reach significance, while for a group of readers, differences were substantial enough to reach significance. Though we included readers of varying experience, effects were too small to prove the effect of readers' experience. This dependence on study composition and reader selection also explains the discordance in results between previous publications.

Because the effect of CAD was beneficial only for a small subgroup of patients with segmental and subsegmental emboli, we could not prove a significant increase in the area under the ROC curve for the complete patient group. These results likely reflect the potential clinical role of CAD for PE: As long as the diagnosis is based on the detection of at least one intraarterial defect in patients who have mostly multiple emboli, the diagnostic effect of CAD on a per-patient basis will be limited to a small number of examinations with isolated and subtle findings and will not be found to significantly improve overall detection performance.

Walsham et al [13] had not been able to show an effect on readers' sensitivity on a per-patient basis when CAD was used. They included 118 consecutive scans and of this number of scans, CAD analysis failed in 18 scans. However, they included a smaller number of readers (one medical student, one resident, and one chest fellow) and only a small number of scans that were positive for PE (n=21). Depending on the severity and conspicuity of PE in these 21 examinations, it is not surprising that they could not find a beneficial effect of CAD.

In a study by Das et al [16], three radiologists evaluated 33 examinations with positive results and 10 examinations with negative results. Because the sensitivity of the readers without CAD was already 100%, they could not show an effect of CAD on a per-patient basis. Furthermore, they excluded scans with poor image quality or underlying lung diseases.

With findings similar to our results, Engelke et al [17] reported an increase in readers' sensitivity from 73%-93% to 77%-97% only for peripheral arteries located on the segmental and subsegmental level with the use of CAD. Their readers were two radiologists and two residents. However, they only evaluated scans that were positive for PE (n=56) and did not analyse the diagnostic effect of CAD on a per-patient basis.

It is known from another study [24] that interobserver variation is largest for diagnosing peripheral emboli. The fact that the increase in sensitivity we found in our study mainly referred to the detection of subsegmental emboli also explains the controversial results of studies so far. Unfortunately, none of the four currently available articles [13, 16-18] dealing with the effect of CAD on reader performance report the number of patients with isolated subsegmental emboli.

We conclude from these findings that CAD is unlikely to have an effect on the detection of obvious or central emboli but, rather, assists readers in detecting peripheral emboli. However, the clinical importance of isolated subsegmental emboli and the need to treat them is still controversial. It is generally accepted thus far that they are mainly important in patients with cardiopulmonary restrictions, recurrent PE or coexisting acute deep venous thrombosis [25].

It is also conceivable that the clinical effect of CAD lies more in reassuring readers than in influencing their diagnosis: There was a significant increase in readers' confidence scores when they used CAD, though the effect of CAD was not the same for all readers. This is noteworthy because this increase did not go along with a significant decrease in specificity or in confidence caused by FP candidate lesions. Not surprisingly, image quality has an important effect on readers' confidence. For all readers, the subjectively scored image quality was significantly correlated with their confidence.

The specificity did not decrease with the use of CAD. On average, CAD identified 4.9 FP candidate lesions per scan (median, three; range, 0 to 42). However, in 8 of 948 (158 cases for six readers) viewed negative scans, readers were misled by FP CAD candidate lesions and did accept one or more candidate lesions as being a PE. We conclude from this result that most of the FP CAD candidate lesions were easy to sort out. Nevertheless, though the percentage of patients that were falsely classified as having PE on the basis of FP CAD candidate lesions remained small, amounting to less than 1% of all examinations, this is an important finding. The risks of treatment without the correct indication should be taken into account on an individual basis.

On average, reading time without CAD was 75 seconds per reader (range, 43-105 seconds) and was extended by 22 seconds per reader (range, 13-31 seconds) with CAD. These reading times were purely for the assessment of

the presence of PE without paying attention to other thoracic abnormalities. For those cases in which readers made a change in the correct diagnosis after using CAD, the mean extra time invested was 62 seconds per scan (range, 18-158 seconds). We found a significant correlation between the number of FP CAD candidate lesions and the additional reading time with the use of CAD. It seems, therefore, that the number of FP CAD candidate lesions should be substantially reduced before CAD is introduced into clinical application.

Our study had the following limitations. Our reference standard was established by two independent readers, with a third reader in case of discordant findings. This reflects clinical reality, where CT has become the standard of reference for PE diagnosis. It is, however, noteworthy, given the fact that the advantage of CAD was seen for subtle emboli, for which it is known that interobserver agreement is the lowest and establishment of a reference standard is the most difficult. Though our study group included more than 200 patients and thus exceeded the size of all previous studies, the number of scans positive for PE was too small to statistically assess the correlation between the performance increase with CAD and reader experience. The fact that we did not find a correlation does not mean that no correlation exists.

Our study was designed to approximate clinical conditions. We did not ask the readers to document lesion locations that served as the basis for their diagnosis. Therefore, we cannot specify the interaction between reader and CAD on a per-lesion basis, only on a per-patient basis. Though this approach does not exclude the possibility that a diagnosis was based on a pseudolesion, we assume this effect to be random and small in a group of more than 200 examinations.

Reading times both without and with CAD were relatively short. This factor was so because we did not ask the readers to localise all emboli, but purely to assess the presence of any emboli to establish the diagnosis of PE . This request was made to approximate the clinical situation of dichotomous decision making as closely as possible. In addition, no attention was paid to the presence of alternative thoracic abnormalities. The true effect of CAD on reading time and patient management needs to be investigated in a prospective clinical study.

In summary, we conclude from our results that CAD has the potential to increase reader sensitivity for the detection of segmental and subsegmental pulmonary emboli and to strengthen reader confidence for the diagnosis of PE without significant loss of specificity. For clinical application of CAD, the number of false positive candidate lesions has to be further decreased.

Reference List

1. Chan HP, Hadjiiski L, Zhou C et al. Computer-aided diagnosis of lung cancer and pulmonary embolism in computed tomography-a review. Acad Radiol 2008; 15(5):535-55.
2. White CS, Pugatch R, Koonce T et al. Lung nodule CAD software as a second reader: a multicenter study. Acad Radiol 2008; 15(3):326-33.
3. Malich A, Fischer DR, Bottcher J. CAD for mammography: the technique, results, current role and further developments. Eur Radiol 2006; 16(7):1449-60.
4. Ghaye B, Remy J, Remy-Jardin M. Non-traumatic thoracic emergencies: CT diagnosis of acute pulmonary embolism: the first 10 years. Eur Radiol 2002; 12(8):1886-905.
5. Remy-Jardin M, Pistolesi M, Goodman LR et al. Management of suspected acute pulmonary embolism in the era of CT angiography: a statement from the Fleischner Society. Radiology 2007; 245(2):315-29.
6. Schoepf UJ, Costello P. CT angiography for diagnosis of pulmonary embolism: state of the art. Radiology 2004; 230(2):329-37.
7. 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(6):651-8.
8. Schoepf UJ, Schneider AC, Das M et al. Pulmonary embolism: computer-aided detection at multidetector row spiral computed tomography. J Thorac Imaging 2007; 22(4):319-23.
9. Zhou C, Chan HP, Patel S et al. Preliminary investigation of computer-aided detection of pulmonary embolism in three-dimensional computed tomography pulmonary angiography images. Acad Radiol 2005; 12(6):782-92.
10. Dewailly M, Remy-Jardin M, Duhamel A et al. Computer-aided detection of acute pulmonary embolism with 64-slice multi-detector row computed tomography: impact of the scanning conditions and overall image quality in the detection of peripheral clots. J Comput Assist Tomogr 2010; 34(1):23-30.
11. Maizlin ZV, Vos PM, Godoy MC et al. Computer-aided detection of pulmonary embolism on CT angiography: initial experience. J Thorac Imaging 2007; 22(4):324-9.
12. Masutani Y, MacMahon H, Doi K. Computerized detection of pulmonary embolism in spiral CT angiography based on volumetric image analysis. IEEE Trans Med Imaging 2002; 21(12):1517-23.
13. Walsham AC, Roberts HC, Kashani HM et al. The use of computer-aided detection for the assessment of pulmonary arterial filling defects at computed tomographic angiography. J Comput Assist Tomogr 2008; 32(6):913-8.
14. 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 2009; 20(4):801-6.

15. Wittenberg R, Peters JF, Sonnemans JJ et al. Impact of image quality on the performance of computer-aided detection of pulmonary embolism. Am J Roentgenol 2011; 196(1):95-101.
16. Das M, Muhlenbruch G, Helm A et al. Computer-aided detection of pulmonary embolism: Influence on radiologists‘ detection performance with respect to vessel segments. Eur Radiol 2008; 18:1350-5.
17. Engelke C, Schmidt S, Bakai A et al. Computer-assisted detection of pulmonary embolism: performance evaluation in consensus with experienced and inexperienced chest radiologists. Eur Radiol 2008; 18(2):298-307.
18. Engelke C, Schmidt S, Auer F et al. Does computer-assisted detection of pulmonary emboli enhance severity assessment and risk stratification in acute pulmonary embolism? Clin Radiol 2010; 65(2):137-44.
19. Buelow T, Wiemker R, Blaffert T et al. Automatic extraction of the pulmonary artery tree from multi-slice CT data. Spie 2005; 5746:730-40.
20. Bouma H, Sonnemans JJ, Vilanova A et al. Automatic detection of pulmonary embolism in CTA images. IEEE Trans Med Imaging 2009;28(8):1223-30.
21. Boyden E. Segmental anatomy of the lungs. New York, NY: McGraw-Hill, 1955.
22. Jackson C, Huber J. Correlated applied anatomy of the bronchial tree and lungs with a system of nomenclature. Dis Chest 1943*; 9:*319 *-326.*
23. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983 September;148(3):839-43.
24. Brunot S, Corneloup O, Latrabe V et al. Reproducibility of multi-detector spiral computed tomography in detection of sub-segmental acute pulmonary embolism. Eur Radiol 2005; 15(10):2057-63.
25. Goodman LR. Small pulmonary emboli: what do we know? Radiology 2005; 234(3):654-8.

Chapter 7

Wittenberg R
Peters JF
van den Berk IAH
Freling NJM
Lely R
de Hoop BJ
Horsthuis K
Ravesloot CJ
Weber M
Prokop M
Schaefer-Prokop CM

Impact on diagnostic performance and reading time of a computer-assisted detection algorithm for the detection of acute pulmonary embolism used as a concurrent reader

Abstract

Purpose

To compare the diagnostic performance and time efficiency of computer-assisted detection (CAD) used as a concurrent reader with reading without CAD for the detection of acute pulmonary embolism (PE) in CT pulmonary angiography (CTPA).

Material and Methods

In this institutional review board-approved retrospective study, six observers of varying experience evaluated 158 negative and 38 positive CTPA (mean age 60; 115 women) without and with CAD as a concurrent reader with an interval of at least 6 weeks. Readers were asked to annotate their diagnosis, confidence score and reading time. Results were compared with an independent standard established by two readers and a third chest radiologist in case of discordant findings.

Results

Using logistic regression for repeated measurements, we found a significant increase in readers' sensitivity ($p<0.001$) without loss of specificity ($p=0.855$), although the effect depended on the reader. A two way ANOVA showed a significant decrease in reading time and an increase in readers' confidence score when CAD was used as a concurrent reader.

Conclusion

CAD as a concurrent reader has the potential to increase readers' sensitivity and confidence with a decrease in reading time without loss of specificity. The effect of CAD, however, significantly differs per reader.

Introduction

Pulmonary embolism (PE) is associated with significant morbidity and mortality. Treatment with anti-clotting medication is highly effective but diagnosis is missed in two third of the cases [1-3]. CT pulmonary angiography (CTPA) has emerged as an effective method for clinical diagnosis of PE, but interpreting a CTPA is complicated by the size of the image date set, the various PE look-alikes and also human factors such as attention span [4, 5]. Computer-assisted detection (CAD) algorithms are therefore under evaluation with respect to their impact on diagnostic performance and reading time.

CAD algorithms are developed to be used as a second reader, assuming that the radiologist first completes a full interpretation of the scan unassisted by CAD before having the CAD candidates available for a "second look". CAD candidates may be added to the interpretation if considered true positive; whether removal of lesions seen by the radiologist but not confirmed by CAD is allowed remains under debate. That way it is ensured that the use of CAD potentially can lead to an increase of sensitivity but at the expense of specificity and unavoidably at the price of prolonged reading time. Furthermore, knowing that the results of the CAD analysis will become available immediately after their initial interpretations, readers may become less conscious. A potentially more time efficient option may be concurrent reading, which allows for use of CAD simultaneously with the readers initial data evaluation.

Previous studies have shown that readers' performance for the detection of PE can be improved using CAD as second reader [6-8]. While the effect of concurrent reading on the detection of PE is not known, studies of CAD algorithms as a concurrent reader for other applications (e.g. mammography, lung nodules or virtual colonoscopy) have already shown a decrease in reading time [9-11].

The purpose of this study was to compare the diagnostic performance and time efficiency of CAD used as a concurrent reader with reading without CAD for the detection of acute PE in CTPA.

Materials and Methods

Study design

In this study we included 200 consecutive CTPA studies performed in a university hospital between January and October 2008. All patients had been referred to the radiology department for suspected PE in clinical routine. All scans were acquired on a 64-slice scanner (Somatom 64, Siemens Medical Solutions, Forchheim, Germany) (Table 1).

Table 1. Computed Tomography protocol on the 64-detector row scanner.

	64-detector row CT
Scan direction	cranio-caudal
kVp	100
mAs	135
Collimation (mm)	24 x 1.2
Rotation time (s)	0.37
Pitch	0.75
Contrast volume (ml)	40-70[A]
Flow (ml/s)	6
NaCl chaser (ml)	30
Delay (s)	5-12
Reconstructed slice thickness (mm)	1.5
Reconstruction-interval (mm)	1.0

[A] Weight adapted

Four CTPA were excluded: in two scans image quality was so strongly hampered by artefacts that diagnosis of PE was impossible. This was due to non-elevated arms in one case and to metallic implants in the spine in the other case. In two scans only findings of chronic PE were seen. The study group therefore consisted of 196 patients: 81 males and 115 females with a mean age of 60 years (median 63, range 20-85).

The reference standard with respect to the location of emboli was established by independent evaluation of the data by a resident specially trained in reading PE studies (> 400 exams, RW) and by a chest radiologist (> 15 year experience, CMS). In case of discordant findings between readers or between readers and the original report, a third experienced chest radiologist (> 15 years experience, MP) was consulted as an arbiter. This reference standard was used to determine the false positive or true positive character of the CAD findings. 19% (38/196) of the studies were positive for PE according to the reference standard and 81% (158/196) negative.

This retrospective data analysis had been approved by the local ethic committee and all scans were anonymised.

Reading methodology

The CAD algorithm (Philips Healthcare, Best, The Netherlands) analyses the data set in the background and processing time therefore has no impact on reading time. After segmentation of the lung and identification of the vascular

structures, the algorithm detects contrast differences within the vascular tree. The algorithm is also trained to detect complete vessel occlusion or contrast-to-tissue transitions but does not differentiate between arteries and veins. Candidate lesions are indicated by coloured circles that are presented on demand.

Six readers with different levels of experience (R1 = 1st year resident, R2 = 3rd year resident, R3 = 3rd year resident, R4 = radiologist with 2 years of experience, R5 = radiologist with 2,5 years of experience, R6 = radiologist with 16 years of experience) independently read all 196 scans blinded to the reference standard. Eight cases, that were not part of the actual study, were used to make them familiar with the use of the software.

Each case was read twice with at least 6 weeks in between. All scans were consecutively divided into 4 subgroups A to D with 50, 50, 50 and 48 cases, respectively. For each case the readers were asked to made their diagnosis using a confidence scale ranging from 1 = "insecure" to 5 = "secure" for the presence or absence of PE, respectively (Table 2). In session 1 the readers read two groups without CAD and 2 groups with CAD as a concurrent reader. In session 2 the readers reread all scans with the alternative reading method. The order of 4 subsets was different for all readers. It was assured that none of the cases were read twice within a single reading session. Because in clinical routine diagnosis is based on a yes or no decision, readers were not asked to document the location of each embolus.

In detail the two reading methods were as follows:

1) The readers evaluated the scan unaware of the CAD results and recorded their diagnosis on a per-patient basis, their confidence score and the reading time without CAD.

2) When CAD was applied as a concurrent reader, the previously stored CAD findings were immediately displayed to the reader during the first read. Diagnosis, confidence and reading time with CAD were recorded.

Statistical analysis

All statistics were done on a per-patient basis and per reader. Sensitivities and specificities without and with CAD as a concurrent reader were compared using logistic regression for repeated measurements, which corrects for multiple readers evaluating the same cases. The terms used in our model were: method (without and with CAD as a concurrent reader) and the interaction between readers and the method. Sensitivities without and with CAD were also compared on an individual reader basis using McNemar tests.

A receiver operating characteristic (ROC) was used for comparison of observer performance for the detection of PE with and without CAD. The areas under curve (A_z) were compared using MedCalc for Windows, version 11.2.1.0 (Med Calc Software, Mariakerke, Belgium) [12].

Table 2. Confidence score.

Score	Description
1	Uncertain if PE is (not) present
2	PE possibly (not) present
3	PE probably (not) present
4	PE very likely (not) present
5	PE definitely (not) present

A two-way ANOVA for repeated measurements was used to compare reading times without CAD and with CAD as concurrent reader. Subsequently, 6 paired T- tests were used to show the exact difference in reading time per reader. To analyse the difference between the average reading time for radiologists compared to the residents also a paired T-test was used. A two-way ANOVA for repeated measurements was used to compare readers' confidence scores without CAD and with CAD.

Results

CAD performance:

The stand alone performance of CAD yielded on a per-patient basis a sensitivity of 95% (36/38) and a specificity of 17% (27/158). CAD marked on average 3.8 FP findings per scan (median 3, range 0-22).

Sensitivity and specificity:

Readers' sensitivity without CAD compared to with CAD used as a concurrent reader changed from 95% (36/38) to 100% (38/38) for R1, 97% (37/38) to 100% (38/38) for R2, 100% (38/38) to 95% (36/38) for R3, 89% (34/38) to 95% (36/38) for R4, 92% (35/38) to 97% (37/38) for R5 and 68% (26/38) to 76% (29/38) for R6, respectively (Figure 1).

Logistic regression for repeated measurements taking all reader data into account revealed a significant increase in readers' sensitivity using CAD ($p<0.001$ using the term CAD), although the effect of CAD was different for all readers ($p<0.001$ using the term reader versus method). On an individual basis, using 6 McNemar tests, we could not show a significant in- or decrease in reader's sensitivity using CAD as a concurrent reader ($p>0.5$). In 14 scans (10 different patients), emboli were missed by the six readers while using CAD as a concurrent reader, though CAD had marked 10 out of this 14 patients as true positive

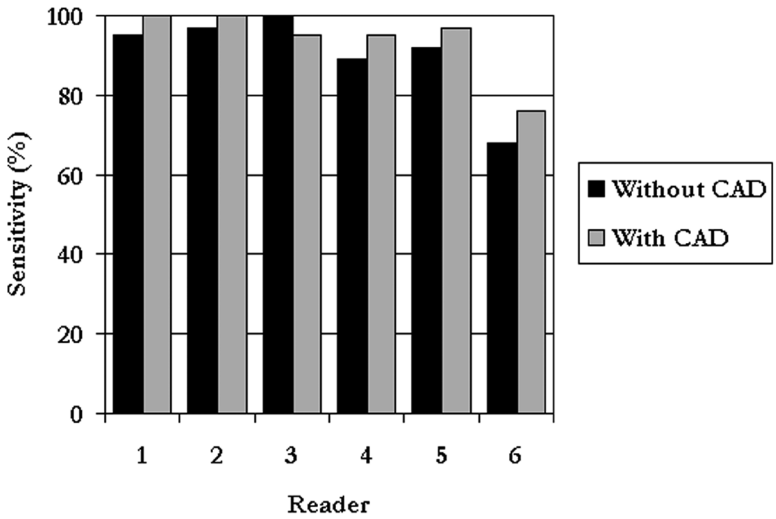

Figure 1. Sensitivity for the six readers.

(71%). In 8 patients the most proximal embolus was at the segmental level and in 2 patients at the lobar level.

Readers specificity without CAD compared to with CAD as a concurrent reader changed from 94% (149/158) to 92% (145/158) for R1, 98% (155 /158) to 99% (156 /158) for R2, 90% (142 /158) to 92% (146/158) for R3, 96% (151/158) to 98% (154 /158) for R4, 95% (150/158) to 90% (142/158) for R5, 95% (150/158) to 97% (154/158) for R6, respectively (Figure 2).

Logistic regression for repeated measurements taking all reader data into account revealed no significant differences in readers' specificity using CAD (p=0.855 using the term CAD), although the effect of CAD was different for all readers (p=0.030 using the term reader versus method). On an individual basis using 6 McNemar tests, a small but significant decrease in the specificity of reader 5 was found (p=0.021).

ROC curves

Only for reader 4 a significant increase of the A_z-value was found. In detail, the A_z-values without CAD compared to with CAD as a concurrent reader amounted to 0.981 and 0.991 (p=0.343) for R1, to 0.979 and 0.996 (p=0.094) for R2, to 0.993 and 0.992 (p=0.594) for R3, to 0.950 and 0.974 (p=0.027) for R4, to 0.971 and 0.990 (p=0.731) for R5, and to 0.928 and 0.914 (p=1.00) for R6, respectively.

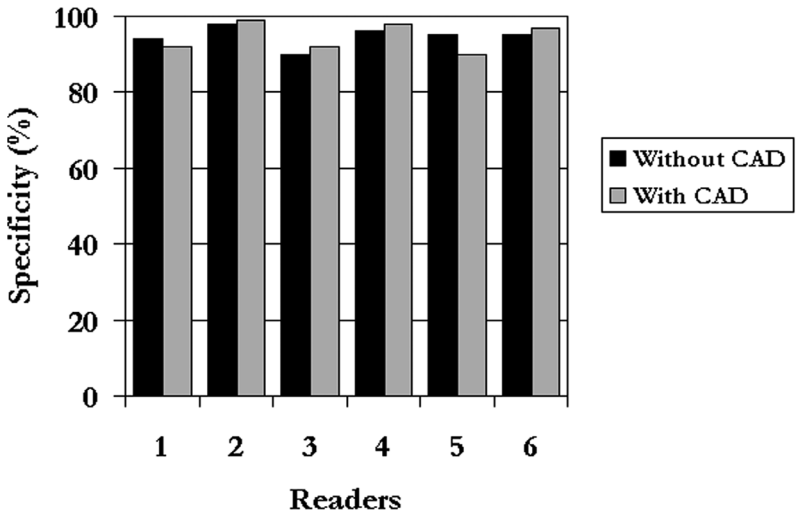

Figure 2. Specificity for the six readers.

Reading time

Using a two-way ANOVA, we found that the average reading time per reader with CAD as a concurrent reader compared to reading without CAD significantly decreased ($p<0.001$), although the effect of CAD depended on the reader ($p<0.001$). Post hoc analysis using 6 paired T-tests showed that for 2 residents and 1 radiologist the reading time significantly decreased, for 1 resident and 1 radiologist the reading time did not differ and that for 1 radiologist the reading time significantly increased. Reading times without CAD and with CAD as a concurrent reader amounted to 208s and 196s for R1 ($p=0.15$), 144s and 118s for R2 ($p=0.003$), 170s and 120s ($p<0.001$) for R3, 24s and 30s ($p<0.001$) for R4, 69s and 66s ($p=0.0323$) for R5 and 32s and 17s ($p<0.001$) for R6 respectively (Figure 3).

The average reading time for residents was significantly higher than for the radiologists for both reading without and with CAD. The average reading time was 174 versus 42 seconds (76%, $p<0.001$) without CAD and 145 versus 37 seconds (74%, $p<0.001$) with CAD as concurrent reader, respectively. The reduction in reading time amounted to 17% and 12% for respectively the residents and radiologists.

Confidence score

Readers' confidence score using a 5-point scale differed between 3.5 and 4.7 for reading without CAD and between 3.9 and 4.7 for reading with CAD as a concurrent reader. We found that readers' confidence significantly increased

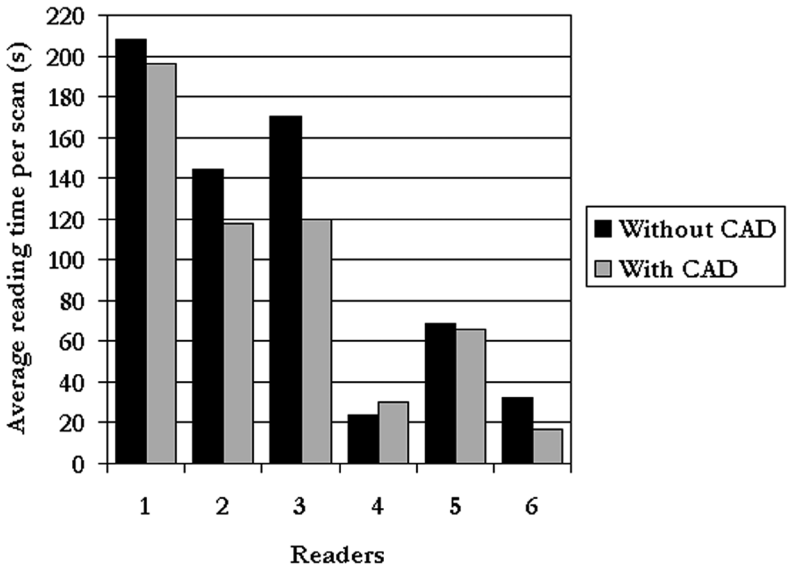

Figure 3. Average reading time per reader without CAD compared to reading with CAD as concurrent reader.

using CAD as a concurrent reader ($p<0.001$), although the effect of CAD on their confidence differed per reader ($p<0.001$).

Discussion

CAD can be integrated in the clinical workflow as a second or a concurrent reader. All articles that have been published on the effect of a CAD prototype on reader performance for the detection of acute PE applied CAD as a second reader [6-8, 13, 14]. Three of them reported a significant increase in readers' sensitivity with CAD[6-8], but none of the five studies reported the effect on reading time.

There is a fear in the radiological community that – if CAD is used as a concurrent reader - readers may be less critical and rely in their decision too much on the CAD candidates with the risk to accept false positive CAD candidates or miss true lesions not indicated by CAD. That is the underlying reason why those CAD algorithms that were FDA approved, are thought to be used as a second reader irrespective for which diagnostic task [15]. Besides a very early type of CAD, none of the CAD prototypes currently under evaluation for the detection of acute PE are FDA approved. Thus experience so far is limited to study conditions. Since previous studies indicated that the diagnostic gain

provided by CAD for the detection of acute PE might be relatively small and refers to only a small subgroup of patients with subtle peripheral emboli [16-18], we suspected that the strength of CAD might rather lie in reducing the reading time than in improving diagnostic performance. A reduction in reading time could be achieved by using CAD concurrently that, however, would be only acceptable if no loss in diagnostic accuracy would be the consequence.

In this study we found a significant increase in sensitivity with CAD as a concurrent reader using logistic regression and taking the whole group of readers into account. The effect of CAD was significantly different for the readers. On individual basis using McNemar tests, the increase of sensitivity did not reach statistical significance. This is likely due to the fact that the basic performance without CAD was already quite high exceeding 90% in 5 out of 6 readers. Besides the number of patients with PE was – as in clinical practice –much smaller than the number of patients without PE. Furthermore among the group of positive patients only few patients suffer from subtle PE. Thus, a very large study group would be necessary to proof an increase of sensitivity for an individual reader. Although the difference was not significant, it should be noted that in one reader (the 3rd year resident) the sensitivity decreased from 100% to 95% using CAD, which in detail referred to the diagnosis in 2 patients. Taking sensitivity and specificity simultaneously into account using ROC analysis, we found that the area under the curve significantly increased only for one radiologist and did not change for the remaining 5.

Taylor et al. [11] had tested CAD as a concurrent reader for the detection of polyps in CT colonography and found that 68% of the polyps that had been missed by readers during concurrent reading had been correctly marked by CAD. The authors' conclusion was that readers may tend to speed up analysis too much when CAD candidates are visible. We found similar results: in 10 of the 14 scans (71%) in which readers missed the diagnosis of pulmonary embolism, the intravascular defects had been correctly identified by CAD. These 14 scans were obtained in 10 different patients. Although we already found an increase of sensitivity with CAD, these results indicate that readers did not fully exploit the advantages of CAD.

Zheng et al. [18] had tested CAD as a concurrent reader for the detection of tumors in mammography. They found that the reduction in sensitivity with the use of CAD as a concurrent reader was correlated with the number of FP CAD candidates, meaning that a too high number of FP candidates makes it more difficult to discriminate between true and false positive candidates. In our study we did not observe a significant loss of sensitivity although the standalone specificity of the CAD algorithm was low with a mean number of false positive candidates of 3.8 per scan. Besides, the visual task in CTPA scans

is fundamentally different from mammography or CT colonography meaning that the detection of any thrombi and not necessarily of all thrombi is sufficient for making the diagnosis of PE.
Not only readers' sensitivity can be influenced by a high number of false positive findings, theoretically also readers' specificity can decrease and reading time can increase. However, this was not confirmed in our study. We found a significant decrease of specificity only in one of the 6 readers, while the specificity of the other 5 readers remained unchanged.

We found that the average reading time significantly decreased when CAD was used as a concurrent reader. The reduction amounted to 17% and 12% respectively for residents and radiologists. The effect, however, differed strongly per reader. On an individual basis, we found a significant decrease for 2 residents and 1 radiologist, while for 1 radiologist the reading time even significantly increased. Overall, reading times were relatively short, because readers were asked only to diagnose pulmonary emboli and not to pay attention to accompanying lung diseases. While our results show the potential of CAD to reduce reading time without loss of diagnostic accuracy for a group of readers, the individual differences in readers' behaviour make it necessary to conduct more research in a larger prospective clinical trial. It is not surprising that for both reading, without CAD and with CAD as a concurrent reader, radiologists needed around 75% less time than residents. Based on the whole reader group, our results also indicated a significant increase in readers' confidence when CAD was used as a concurrent reader. On an individual basis, however, differences were relatively small and did not reach significance.

Our study suffers from the following limitations. Theoretically, the order of the two readings could have caused a learning effect. Its impact on results was kept negligibly low by alternating the reading order. We found significant differences for the group of six readers, however, we found only very few significant differences on individual reader basis. This is likely to be related to the too small number of positive scans, although our study group of 200 consecutive CTPA exceeded the size of groups in previous publications.

In summary, CAD used as a concurrent reader has the potential to increase readers' sensitivity and confidence and at the same time decreases reading time without loss of specificity. However, the relatively large differences between readers make the evaluation of CAD as a concurrent reader in a larger trial necessary before it can be generally recommended to use CAD concurrently and not as a second reader for the detection of acute PE.

Reference List

1. British Thoracic Society guidelines for the management of suspected acute pulmonary embolism. Thorax 2003; 58(6):470-83.
2. Goldhaber SZ, Visani L, De Rosa M. Acute pulmonary embolism: clinical outcomes in the International Cooperative Pulmonary Embolism Registry (ICOPER). Lancet 1999; 353(9162):1386-9.
3. Tapson VF. Acute pulmonary embolism. N Engl J Med 2008; 358(10):1037-52.
4. Remy-Jardin M, Pistolesi M, Goodman LR et al. Management of suspected acute pulmonary embolism in the era of CT angiography: a statement from the Fleischner Society. Radiology 2007; 245(2):315-29.
5. Schoepf UJ, Costello P. CT angiography for diagnosis of pulmonary embolism: state of the art. Radiology 2004; 230(2):329-37.
6. Engelke C, Schmidt S, Bakai A et al. Computer-assisted detection of pulmonary embolism: performance evaluation in consensus with experienced and inexperienced chest radiologists. Eur Radiol 2008; 18(2):298-307.
7. Engelke C, Schmidt S, Auer F et al. Does computer-assisted detection of pulmonary emboli enhance severity assessment and risk stratification in acute pulmonary embolism? Clin Radiol 2010; 65(2):137-44.
8. Blackmon KN, Florin C, Bogoni L et al. Computer-aided detection of pulmonary embolism at CT pulmonary angiography: can it improve performance of inexperienced readers? Eur Radiol 2011; 21(6):1214-23.
9. Beyer F, Zierott L, Fallenberg EM et al. Comparison of sensitivity and reading time for the use of computer-aided detection (CAD) of pulmonary nodules at MDCT as concurrent or second reader. Eur Radiol 2007; 17(11):2941-7.
10. Paquerault S, Samuelson FW, Petrick N et al. Investigation of reading mode and relative sensitivity as factors that influence reader performance when using computer-aided detection software. Acad Radiol 2009; 16(9):1095-107.
11. Taylor SA, Charman SC, Lefere P et al. CT colonography: investigation of the optimum reader paradigm by using computer-aided detection software. Radiology 2008; 246(2):463-71.
12. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983; 148(3):839-43.
13. Das M, Muhlenbruch G, Helm A et al. Computer-aided detection of pulmonary embolism: Influence on radiologists‘ detection performance with respect to vessel segments. Eur Radiol 2008; 18:1350-5.
14. Walsham AC, Roberts HC, Kashani HM et al. The use of computer-aided detection for the assessment of pulmonary arterial filling defects at computed tomographic angiography. J Comput Assist Tomogr 2008; 32(6):913-8.
15. Wittenberg R, Bergers FH, Peters JF, et al. Acute pulmonary embolism: Effect of a computer-assisted detection prototype on diagnosis - An observer study. Radiology 2012 Jan; 262(1):305-13.

16. 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(6):651-8.
17. Schoepf UJ, Schneider AC, Das M et al. Pulmonary embolism: computer-aided detection at multidetector row spiral computed tomography. J Thorac Imaging 2007; 22(4):319-23.
18. Zheng B, Swensson RG, Golla S et al. Detection and classification performance levels of mammographic masses under different computer-aided detection cueing environments. Acad Radiol 2004; 11(4):398-406.

Chapter 8

General discussion

General discussion

Pulmonary embolism (PE) is a common and often potentially life threatening disease. Severe morbidity and mortality can be prevented if PE is diagnosed and treated in time. Currently, anticoagulation therapy is based on a "yes" or "no" decision for the presence of PE, not taking into account the extent of the disease. CT pulmonary angiography is the diagnostic method of choice for PE. Computer-assisted detection algorithms are designed to help radiologists in finding PE. In this thesis, a CAD prototype from Philips Healthcare has been tested; results provided a first estimation of its diagnostic potential but also of its weaknesses that require further improvement before introducing the software into clinical application.

In a number of studies, the standalone performance of CAD as well as its impact on observer performance was determined [1-11]. To put the results into perspective, a reference standard for the presence or absence of PE was needed. The definition of such a reference standard, however, represents a methodological challenge because CTPA is considered the most sensitive imaging method for the diagnosis of PE and its results cannot be compared to results of another "superior" diagnostic method. We defined our standard on the consensus of multiple readings that involved readers of different experience levels. All scans were read by the researcher specially trained in reading CTPA studies and a chest radiologist with more than 15 years of experience. In cases of discrepancy between these two readers or between these readers and the original report, a third experienced chest radiologist served as arbiter. This is not an ideal gold standard, especially for the definition of small subsegmental emboli that turned out to be important for the assessment of the effects of CAD on reader behavior. On the other hand, this standard represents the best that can be created under clinical conditions and is used in all reported articles on CAD prototypes for PE.

With at least 240 consecutive CTPA scans per study, our study group exceeded the size of all study groups in previous publications. To approximate clinical conditions, we also did not select examinations to include only those with high image quality. For determining the standalone performance of CAD, we included only CTPA scans obtained during night and weekend shifts. We did so because this group of scans is read at least initially by a heterogeneous group of readers including residents on call and general radiologists with differing experience and may therefore take advantage of the availability of a CAD system. Both, the standalone performance of CAD as well as the observer studies revealed that CAD has the following three potentials:

1) to increase the detection of peripheral emboli.
2) to increase readers' diagnostic confidence.
3) to decrease reading time, if CAD is used concurrently.

Increasing the detection of peripheral emboli

When we tested the standalone performance of CAD, we found the strength of CAD to lie in the detection of small peripheral emboli, which are diagnostically important if no more proximal emboli are present. This was confirmed by the results of our 2 observer studies. When CAD was used as a second reader, readers' sensitivity increased within a range of 2% to 13% without significant loss of specificity. Furthermore, when CAD was used as a concurrent reader, readers' sensitivity remained unchanged and only one reader showed a small significant decrease in specificity. In both studies almost all emboli found with the help of CAD were located at the (sub)segmental level. These results suggests that CAD will only have a potential diagnostic impact in a subgroup of patients with subtle peripheral emboli, but not in patients with rather obvious, central PE. However, the clinical importance of small isolated emboli is still uncertain. It is suggested that small pulmonary emboli are predictors of more severe embolic events and that they are especially relevant in patient groups at risk, for example patients with cardiopulmonary restriction [12]. It was discussed that detection of these small emboli may even lead to overtreatment, considering the potential complications of anticoagulation therapy [13]. In this context, it may even be the case that CAD further aggravates this clinical dilemma. If, in the future, treatment would be stronger individualised and based on thrombus load instead of a yes or no decision, CAD could play an important role. Recently, perfusion imaging of the lung parenchyma became available [14]. In that context the complementary information of complete delineation of all intravascular defects supported by CAD and quantification of resulting perfusion defects may improve diagnostics for PE.

Increasing readers' diagnostic confidence

In our studies, the prevalence of PE was around 20%, but in the literature even smaller numbers were published. Therefore, it is not only important to note that CAD will help readers to find peripheral emboli, but that it also reassures readers in excluding PE. In around 80% of our studies, we reported an average number of FP CAD findings between 3.8 and 6.2. However, if CAD did not find any candidates, it had a negative predictive value of more than 90% that indeed no PE was present. Additionally, we found that readers' confidence significantly increased with CAD. The combination of a high negative predictive value and an increase in readers' confidence helps to securely exclude PE. Unfortunately,

the number of scans without any FP CAD candidate was relatively small and a further decrease of FP CAD candidates is absolutely mandatory if CAD is to be used successfully in clinical practice. We did not observe a significant loss of specificity in our observer studies, indicating that it is quite easy for the readers to discriminate between true and false positive candidates. Because thirty percent of the FP CAD findings were located in veins, implementation of an artery-vein separation within the CAD analysis appears to be an effective step. We also showed that not the scanner type, but mainly the scanning protocol is a determining factor for the standalone performance of CAD. A refinement of the scanning protocol to reduce false positive CAD findings seems to be another aspect that needs further research.

Decreasing reading time

Up until now, all CAD prototypes for the detection of PE reported in the literature were developed to be used as a second reader. Because with CAD as a second reader, CAD findings are made visible after a complete initial assessment by the reader without assistance of CAD, an increase in reading time is inevitable. In this thesis we showed that reading time will significantly decrease if CAD is used as a concurrent reader, meaning it being used from the start of the assessment. More importantly, readers' sensitivity did not decrease and in only one reader did the specificity slightly decrease with this reading method. Because there are no other studies testing CAD as a concurrent reader thus far, these findings need to be tested in a larger prospective clinical trial to investigate the effects of CAD as a concurrent reader on patient outcome. As discussed before in the previous paragraph, the number of FP CAD candidates needs to be reduced not only to increase readers' confidence into the capability of the CAD analysis but also to keep reading times acceptable.

In **summary,** CAD has the potential to aid readers to find patients with isolated segmental and/or subsegmental emboli, to reassure readers in excluding PE and to decrease reading time. These advantages can be further strengthened if CAD will be improved, especially with respect to the number of FP candidates. For that purpose, a discrimination of arteries and veins during the analysis appears to be a first effective step. In the future, calculation of thrombus load may play a role for the improvement of therapy and assessment of prognosis. Before making CAD available for broader application in clinical practice, its impact on readers' decisions needs to be studied in a larger prospective trial.

Reference List

1. Masutani Y, MacMahon H, Doi K. Computerized detection of pulmonary embolism in spiral CT angiography based on volumetric image analysis. IEEE Trans Med Imaging 2002; 21(12):1517-23.
2. Walsham AC, Roberts HC, Kashani HM et al. The use of computer-aided detection for the assessment of pulmonary arterial filling defects at computed tomographic angiography. J Comput Assist Tomogr 2008;32(6):913-8.
3. Zhou C, Chan HP, Patel S et al. Preliminary investigation of computer-aided detection of pulmonary embolism in three-dimensional computed tomography pulmonary angiography images. Acad Radiol 2005;12(6):782-92.
4. Blackmon KN, Florin C, Bogoni L et al. Computer-aided detection of pulmonary embolism at CT pulmonary angiography: can it improve performance of inexperienced readers? Eur Radiol 2011; 21(6):1214-23.
5. Buhmann S, Herzog P, Liang J et al. Clinical evaluation of a computer-aided diagnosis prototype for the detection of pulmonary embolism. Acad Radiol 2007; 14(6):651-8.
6. Das M, Muhlenbruch G, Helm A et al. Computer-aided detection of pulmonary embolism: Influence on radiologists' detection performance with respect to vessel segments. Eur Radiol 2008; 18:1350-5.
7. Dewailly M, Remy-Jardin M, Duhamel A et al. Computer-aided detection of acute pulmonary embolism with 64-slice multi-detector row computed tomography: impact of the scanning conditions and overall image quality in the detection of peripheral clots. J Comput Assist Tomogr 2010; 34(1):23-30.
8. Engelke C, Schmidt S, Bakai A, Auer F et al. Computer-assisted detection of pulmonary embolism: performance evaluation in consensus with experienced and inexperienced chest radiologists. Eur Radiol 2008; 18(2):298-307.
9. Engelke C, Schmidt S, Auer F et al. Does computer-assisted detection of pulmonary emboli enhance severity assessment and risk stratification in acute pulmonary embolism? Clin Radiol 2010; 65(2):137-44.
10. Maizlin ZV, Vos PM, Godoy MC et al. Computer-aided detection of pulmonary embolism on CT angiography: initial experience. J Thorac Imaging 2007; 22(4):324-9.
11. Schoepf UJ, Schneider AC, Das M et al. Pulmonary embolism: computer-aided detection at multidetector row spiral computed tomography. J Thorac Imaging 2007; 22(4):319-23.
12. Schoepf UJ, Costello P. CT angiography for diagnosis of pulmonary embolism: state of the art. Radiology 2004; 230(2):329-37.
13. Goodman LR. Small pulmonary emboli: what do we know? Radiology 2005; 234(3):654-8.
14. Thieme SF, Johnson TR, Reiser MF et al. Dual-energy lung perfusion computed tomography: a novel pulmonary functional imaging method. Semin Ultrasound CT MR 2010; 31(4):301-8.

Chapter 9

Summary & Samenvatting

Summary

In this thesis, we evaluated the performance of a computer-assisted detection (CAD) prototype for pulmonary embolism (PE). The first part outlines the most recent technical advances of CT pulmonary angiography (CTPA) for the detection of PE with special emphasis on aspects of image quality. The evaluation of a CAD prototype is described in part 2 and 3: part 2 is dealing with the standalone performance of the software and the impact of image quality on it; part 3 evaluates how CAD actually influences reader performance.

Part 1: Update on detecting pulmonary embolism using CT angiography

Since the introduction of CT scanners with multiple detector rows in 1992, the quality of CT pulmonary angiography (CTPA) and its diagnostic performance have significantly improved. While the first scanners provided the technical break-through in terms of continuous volume data acquisition, it suffered from a relatively low spatial resolution in the patient length axis (z-axis) and still needed more than 24 seconds for scanning the whole chest. Most modern CT scanners, however, with up to 256 (2 x 128) and 320 detector rows allow for scanning the thorax in less than 3 seconds with a submillimeter isotropic resolution. Good quality images, meaning with homogeneously high intravascular contrast and no or minimal movement artefacts can be obtained even in very dyspneic patients. Accurate analysis of small subsegmental arteries and reliable detection of very small emboli are possible.

However, also with the most modern scanners, image quality is the result of data acquisition, contrast injection and patient behavior. The extremely short period of data acquisition requires an adequate protocol of contrast injection. Patient movement or a patient's Valsalva maneuver at the wrong time can destroy the quality of the whole scan.

The high resolution of the images and the multitude of information that has to be visually assessed were the motivation for the development of elaborate processing algorithms and analytic software tools. All these digital tools are designed to support the diagnostic evaluation by the radiologist, to compensate for perception mistakes, to shorten the reading time and to ease the visual analysis.

Chapter 2 summarises the latest technical development of CT scanners, the new challenges for protocol design associated with these new techniques and the options offered by new processing and evaluation tools.

Part 2: Standalone performance of a computer-assisted detection prototype for the detection of pulmonary embolism

In **chapter 3**, we assessed the standalone performance of CAD in a group of 278 CTPA examinations that had been consecutively obtained during night shifts and weekends. According to the reports, 68 patients were found to be positive for PE. Our results indicated that CAD found emboli in 7 additional patients that had been originally reported as negative for the presence of PE. In all 7 patients, PE were located relatively peripheral on the subsegmental or segmental level. In total, the study group contained 14 patients with isolated subsegmental emboli, meaning with no emboli further proximally, and CAD correctly identified 13 of these 14 patients. These findings indicate that the strength of CAD lies in alerting the radiologist to the presence of small peripheral emboli that are most subtle and most difficult to see. Transferring these results into a clinical situation means that CAD will have a potential diagnostic impact only in a subgroup of patients with subtle peripheral emboli. Since the diagnosis of PE does only require the detection of *any* emboli and not the detection of *all* emboli, CAD is unlikely to provide additional diagnostic value in patients with rather obvious and more conspicuous emboli.

The prevalence of PE in CTPA scans ranges from 10 to 35%, thus the majority of CTPA examinations are made to exclude PE. If there were no CAD candidates at all, the negative predictive value of such a CAD analysis exceeded 90% that indeed no PE was present. However, in most negative scans, CAD showed false positive findings (on average 4.7 per scan) with an overall specificity of 21% underlining the need to further decrease the number of false positive candidates.

It is likely that the quality of the CTPA scan will influence the performance of the CAD software. In **chapter 4,** we therefore focused on the impact of various aspects of image quality on the performance of CAD. We found a strong association between the number of false positive CAD findings and the overall quality of the scan, the vascular enhancement, the presence of motion artefacts and underlying lung disease and the severity of noise. We also showed that the majority of the FP CAD candidates were located in veins or in arteries surrounded by consolidations.

Both studies described in chapters 3 and 4, included 100 patients that had been scanned with a 64-slice scanner and 178 patients that had been scanned with a 16-slice scanner. The 64 slice scanner has a potential advantage over the 16-slice scanner based on its higher spatial resolution in the z-axis and the faster scan acquisition resulting in less pulsation or breathing artefacts. However, our results indicated a significantly lower number of FP CAD findings for the 16-slice CTPA scan than for the 64-slice scans: 3.8 versus 6.3 respectively.

Furthermore, in the 16-slice scans, noise was significantly lower and contrast enhancement and overall quality significantly higher than in the 64-slice scans. These results underline the importance of the scan protocol for the image quality and subsequently the performance of CAD, overruling the technical advantage of the scanner type.

In clinical practice, various scanner types and scan protocols are used for the diagnosis of PE. The purpose of the study described in **chapter 5** was therefore to compare the standalone performance of CAD in CTPA examinations obtained with three different 64-slice scanner types (Philips, GE and Siemens) in 3 different hospitals. Although the CTPA examinations of the three hospitals significantly differed with respect to image noise, vascular enhancement and the number of FP CAD candidates, the sensitivity and specificity on a per patient basis were comparable, ranging from 92% to 100% and 13% to 18%, respectively. When we corrected our results for the scan quality, scanner type did not affect the number of FP CAD findings per scan anymore. We concluded from these findings that the variation of image quality within a certain range, inevitably present when using different scanner types and protocols, does not lead to unacceptable performance differences of the CAD prototype with respect to the diagnosis of PE per patient.

Part 3: Impact of a computer-assisted detection prototype used as a second or a concurrent reader on readers with varying experience

Whether CAD can improve the detection of PE in a clinical setting depends not only on the standalone performance of the CAD software, but also on the effect of CAD on readers' decisions. To investigate this effect, we asked four residents and two radiologists to assess 209 CTPA scans first without CAD followed immediately by a second assessment with the availability of CAD candidates within a single reading session. The methodology and results have been described in **chapter 6**.

CAD helped the readers to detect PE in 15 patients. In all patients, emboli were located only at the (sub)segmental level and had been missed during the first unassisted evaluation. With CAD, readers' sensitivity showed a significant improvement within a range of 2% to 13% without significant loss of specificity. Similarly as seen in part 2, we found the greatest impact of CAD on the detection of small segmental and subsegmental emboli.

We also found that readers' confidence with CAD significantly improved, although the magnitude of the effect depended on the reader. These results suggest that CAD is likely to have a larger effect on inexperienced than experienced readers. In order to show this, statistically larger reader and patient groups are required.

The increase of sensitivity and confidence came with an increase of reading time, which is undesirable in clinical practice The increase in reading time is inevitable if CAD is applied as a second reader because it requires a second scan evaluation after a first unassisted image interpretation. When CAD is used as a concurrent reader, meaning that the CAD candidates have already been identified during the first image analysis, a decrease in reading time can be potentially accomplished. In **chapter 7**, we describe a study that compared the performance of three residents and three radiologists reading 196 CTPA examinations with CAD as a concurrent reader versus reading without CAD. When using CAD concurrently, readers' sensitivity improved or remained at the same level and readers' confidence increased compared to the readings without CAD with a significant shorter reading time.

Samenvatting

In dit proefschrift hebben we het functioneren van een 'computer-assisted detection' (CAD) prototype voor het diagnosticeren van longembolieën onderzocht. In het eerste deel van dit proefschrift worden de meest recente technische ontwikkelingen van de CT pulmonaire angiografie (CTPA) voor de detectie van longembolieën besproken, waarbij de nadruk ligt op de beeldkwaliteit. De evaluatie van een CAD prototype wordt beschreven in deel 2 en 3, waarbij in deel 2 het zelfstandig functioneren van de software en de impact van de beeldkwaliteit op deze software wordt besproken. In deel 3 wordt onderzocht hoe CAD de prestaties van lezers beïnvloedt.

Deel 1: Het detecteren van longembolieën met behulp van de CT-angiografie: een update.

Sinds de introductie van de multidetector CT in 1992, is de kwaliteit en het diagnostisch functioneren van de CTPA aanzienlijk verbeterd. Terwijl de eerste scanners een technische doorbraak gaven wat betreft de continue data-acquisitie, leverden deze scans een relatief lage spatiële resolutie in de lengte-as (z-as) op en was er nog steeds meer dan 24 seconden nodig voor het scannen van de hele thorax. De meeste moderne CT scanners met maximaal 256 (2 x 128) en 320 detector rijen, kunnen echter de thorax scannen in minder dan 3 seconden met een submillimeter isotrope resolutie. Beelden van goede kwaliteit, dat wil zeggen met goed homogeen intravasculair contrast en geen of minimale bewegings artefacten, kunnen zelfs bij zeer dyspnoïsche patiënten worden verkregen. Ook is nauwkeurige analyse en betrouwbare detectie van kleine subsegmentele arteriën mogelijk.

Echter, ook met de meest moderne scanners, is de beeldkwaliteit het resultaat van data-acquisitie, contrast injectie en het meewerken van de patiënt. Omdat de data-acquisitie in een erg korte tijd plaatsvindt is een adequaat contrast protocol vereist. Beweging van de patiënt of een Valsalva manoeuvre op het verkeerde moment kan de kwaliteit van de scan negatief beïnvloeden.

De hoge resolutie van de beelden en de hoeveelheid aan informatie die visueel beoordeeld moet worden, waren de motivatie voor het ontwikkelen van algoritmes en analytische software. Al deze digitale hulpmiddelen zijn ontworpen om de diagnostische beoordeling door de radioloog te ondersteunen, om te compenseren voor de perceptie fouten, om de leestijd te verkorten en om de visuele analyse te vergemakkelijken.

Hoofdstuk 2 geeft een overzicht van de nieuwste technische ontwikkelingen van CT scanners, de uitdagingen voor het ontwikkelen van een scan protocol en de mogelijkheden van nieuwe processing en evaluatie software.

Deel 2: Het zelfstandig functioneren van een CAD prototype voor de detectie van longembolieën.

In **hoofdstuk 3** onderzochten we het zelfstandig functioneren van het CAD in een groep van 278 CTPA scans die consecutief waren verkregen tijdens nacht- en weekenddiensten. Volgens de verslagen bleken 68 patiënten longembolieën te hebben. Onze resultaten lieten zien dat CAD in zeven extra patiënten, die oorspronkelijk als negatief beoordeeld waren, toch longembolieën vond. Bij alle zeven patiënten bevonden de longembolieën zich perifeer op het subsegmentele of segmentele niveau. In de totale studiegroep waren er 14 patiënten met geïsoleerde subsegmentele longembolieën, dat wil zeggen zonder embolieën proximaal van dit niveau. CAD identificeerde 13 van deze 14 patiënten correct. Deze bevindingen wijzen erop dat de kracht van CAD ligt in het alarmeren van de radioloog op kleine perifere embolieën, die klinisch ook het meest moeilijk te detecteren zijn. In een klinische situatie betekent dit dat CAD alleen een potentieel diagnostisch effect zal hebben in een subgroep van patiënten. Omdat voor de diagnose van longembolieën slechts de detectie van één embolie vereist is en niet de detectie van alle embolieën, is het onwaarschijnlijk dat CAD aanvullende diagnostische waarde heeft bij patiënten met opvallende embolieën.

Omdat de prevalentie van longembolieën in CTPA scans varieert tussen de 10 en 35%, wordt de meerderheid van de CTPA scans gemaakt om longembolieën uit te sluiten. Als CAD geen verdachte laesies laat zien, is de negatief voorspellende waarde van een dergelijke CAD analyse meer dan 90% dat er inderdaad geen embolieën aanwezig zijn. Echter, in de meeste negatieve scans, laat CAD toch fout positieve laesies zien (gemiddeld 4,7 per scan) waardoor de specificiteit van CAD 21% is. Dit onderstreept de noodzaak om het aantal fout positieve laesies verder te reduceren.

Het is waarschijnlijk dat de kwaliteit van de CTPA scan invloed zal hebben op de prestaties van de CAD software. In **hoofdstuk 4** hebben we ons daarom op de invloed van verschillende aspecten van de beeldkwaliteit op de prestaties van CAD gericht. We vonden een sterke associatie tussen het aantal fout positieve CAD laesies en de algehele kwaliteit van de scan, het contrast aanbod, de aanwezigheid van bewegingsartefacten en onderliggende longziekten en de mate van ruis. We toonden ook aan dat de meerderheid van de fout positieve CAD laesies zich bevonden in venen dan wel in arteriën die omgeven waren door consolidaties.

Voor beide studies beschreven in hoofdstuk 3 en 4, waren 100 patiënten gescand met een 64-slice scanner en 178 patiënten met een 16-slice scanner. De 64-slice scanner heeft een potentieel voordeel ten opzichte van de 16-slice scanner, omdat de spatiële resolutie in de z-richting hoger is en de snellere scan resulteert in minder pulsatie- of ademartefacten. Echter, onze resultaten lieten

een significant lager aantal fout positieve CAD laesies zien voor de 16-slice scans dan voor de 64-slice scans: 3,8 versus 6,3 respectievelijk. Bovendien was in de 16-slice scans de mate van ruis aanzienlijk lager en het contrast aanbod en de algehele kwaliteit aanzienlijk hoger dan in de 64-slice scans. Deze resultaten benadrukken het belang van het scan protocol voor de beeldkwaliteit en dus voor de prestaties van CAD, waarbij het technische voordeel van het scanner type teniet gedaan wordt.

In de kliniek worden diverse soorten scanners en scan protocollen gebruikt voor de diagnose van longembolieën. Het doel van de studie beschreven in **hoofdstuk 5** was dan ook om het zelfstandig functioneren van CAD te vergelijken in CTPA scans die verkregen zijn met drie verschillende 64-slice scanners (Philips, GE en Siemens) in 3 verschillende ziekenhuizen. Hoewel de CTPA scans van de drie ziekenhuizen significant verschilden met betrekking tot de beeldruis, contrastaanbod en het aantal van fout positieve CAD laesies, waren de sensitiviteit en specificiteit op patiënten basis vergelijkbaar. Dit varieerde van 92% tot 100% en van 13% tot 18 %, respectievelijk. Wanneer we onze resultaten corrigeerden voor de scan kwaliteit, had het scanner type geen invloed meer op het aantal fout positieve CAD laesies per scan. We concludeerden uit deze bevindingen dat de variatie van de beeldkwaliteit binnen een bepaald bereik, wat onvermijdelijk aanwezig is bij het gebruik van verschillende scanner types en protocollen, niet leidt tot onaanvaardbare verschillen in de prestaties van het CAD prototype.

Deel 3: Effect van een CAD prototype, gebruikt als 'second of concurrent' reader, op lezers met wisselende ervaring

Of CAD de detectie van longembolieën in een klinische setting kan verbeteren is niet alleen afhankelijk van het zelfstandig functioneren van de CAD software, maar ook van het effect van CAD op lezers. Om dit effect te onderzoeken, vroegen we vier radiologie assistenten in opleiding en twee radiologen om 209 CTPA scans te beoordelen, eerst zonder CAD en onmiddellijk gevolgd door een tweede evaluatie met de beschikbaarheid van de CAD bevindingen. De methodologie en de resultaten van dit onderzoek zijn beschreven in **hoofdstuk 6.**

CAD hielp de lezers om longembolieën te detecteren in 15 patiënten. Bij deze patiënten bevonden de longembolieën zich alleen op het (sub)segmentele niveau en werden gemist tijdens de eerste lezing zonder hulp van CAD. Met CAD toonde de sensitiviteit van de lezers een significante verbetering uiteenlopend van 2% tot 13% zonder een significant verlies van de specificiteit. Net zoals in deel 2, vonden we dat CAD het grootste effect heeft bij de detectie van kleine segmentele en subsegmentele embolieën.

Tevens vonden we dat het zelfvertrouwen van de lezers met behulp van CAD aanzienlijk verbeterde, alhoewel de grootte van het effect afhankelijk was

van de lezer zelf. Deze resultaten suggereren dat CAD waarschijnlijk een groter effect heeft op onervaren dan ervaren lezers. Echter, om dit aan te tonen zijn statistisch groter groepen met lezers en patiënten nodig.

De toename van de sensitiviteit en het zelfvertrouwen van de lezers ging gepaard met een stijging van de leestijd hetgeen ongewenst is in de klinische praktijk. De stijging van de leestijd is onvermijdelijk als CAD wordt toegepast als 'second reader', omdat het CAD pas na een eerste zelfstandig beeldinterpretatie wordt ingezet. Wanneer CAD wordt gebruikt als 'concurrent reader', wat betekent dat de CAD bevindingen al tijdens de eerste evaluatie zichtbaar worden gemaakt, kan mogelijk een afname van de leestijd worden gerealiseerd. In **hoofdstuk 7** gebruikten we 196 CTPA scans om het effect van CAD als 'concurrent reader' op drie radiologie assistenten in opleiding en drie radiologen te vergelijken met het lezen zonder CAD. Uit ons onderzoek kwam naar voren dat bij het gebruik van CAD als 'concurrent reader' de sensitiviteit van lezers verbeterde of op hetzelfde niveau bleef, hun zelfvertrouwen significant toe nam en de leestijd significant korter werd vergeleken met het beoordelen van CTPA scans zonder CAD.

Chapter 10

Publications in international journals

Wittenberg R, Peters JF, van den Berk IAH, Freling NJM, Lely R, de Hoop BJ, Horsthuis K, Ravesloot CJ, Weber M, Prokop M, Schaefer-Prokop CM. Impact on diagnostic performance and reading time of a computer-assisted detection algorithm for the detection of acute PE used as a concurrent reader. *Submitted.*

Wittenberg R, Berger FH, Peters JF, Weber M, van Hoorn F, Beenen LFM, van Doorn MMAC, van Schuppen J, Zijlstra IJAJ, Prokop M, Schaefer-Prokop CM. Acute pulmonary embolism: effect of a computer-assisted detection prototype on diagnosis—An observer study.
Radiology 2012 Jan; 262(1):305-13.

Wittenberg R, Peters JF, Weber M, Lely RJ, Cobben LPJ, Prokop M, Schaefer-Prokop CM.
Stand-alone performance of a computer-assisted detection prototype for detection of acute pulmonary embolism: a multi-institutional comparison.
British J Radiol 2011 Dec (Epub ahead of print).

Wittenberg R, van Vliet JW, Ghaye B, Peters JF, Schaefer-Prokop CM, Coche E. Comparison of automated 4-chamber cardiac views versus axial views for measuring right ventricular enlargement in patients with suspected pulmonary embolism.
Eur J Radiol 2012 Feb; 81(2):218-22.

Wittenberg R, Peters JF, Sonnemans JJ, Bipat S, Prokop M, Schaefer-Prokop CM. Impact of image quality on the performance of a computer-assisted detection prototype for pulmonary embolism.
Am J Roentgenol 2011 Jan; 196(1):95-101.

Wittenberg R, Peters JF, Sonnemans JJ, Prokop M, Schaefer-Prokop CM. Computer-assisted detection of pulmonary embolism: evaluation of pulmonary CT angiographies performed in an on call setting.
Eur Radiol 2010 Apr; 20(4):801-6.

Hartmann IJ, **Wittenberg R**, Schaefer-Prokop CM.
Imaging of acute pulmonary embolism using multi-detector CT angiography: an update on imaging technique and interpretation.
Eur J Radiol 2010 Apr; 74(1):40-9.

v Rijn RR, Boot A, **Wittenberg R**, van der Sluis IM, van den Heuvel-Eibrink MM, Lequin MH, de MuinckKeizer-Schrama SM, van Kuijk C.
Direct X-ray radiogrammetry versus dual-energy X-ray absorptiometry: assessment of bone density in children treated for acute lymphoblastic leukaemia and growth hormone deficiency.
Pediatric Radiology 2006 Mar; 36:227-232.

Conference proceedings - oral presentations

Wittenberg R, Peters JF, van den Berk IAH, Freling NJM, Lely R, de Hoop BJ, Horsthuis K, Ravesloot CJ, Weber M, Prokop M, Schaefer-Prokop CM.
Impact on diagnostic performance and reading time of a computer-aided detection algorithm for the detection of acute PE: second reading versus concurrent reading.
Radiologendagen, 29-30 September 2011, Maastricht, TheNetherlands.

Wittenberg R, Peters JF, van den Berk IAH, Freling NJM, Lely R, de Hoop BJ, Horsthuis K, Ravesloot CJ, Weber M, Prokop M, Schaefer-Prokop CM.
Impact on diagnostic performance and reading time of a computer-aided detection algorithm for the detection of acute PE: second reading versus concurrent reading.
European Congress of Radiology, 3-7 March 2011, Vienna, Austria.

Wittenberg R, van Vliet JW, Ghaye B, Peters JF, Schaefer-Prokop CM, Coche E.
Comparison of automated 4-chamber cardiac views versus axial views for measuring right ventricular enlargement in patients with suspected pulmonary embolism.
Radiologendagen 2010, 16-17 September, Veldhoven, The Netherlands.

Wittenberg R, Berger FH, Peters JF, Weber M, van Hoorn F, Beenen LFM, van Doorn MMAC, van Schuppen J, Zijlstra IJAJ, Prokop M, Schaefer-Prokop CM.
Impact of a computer-assisted detection prototype on the detection of acute pulmonary embolism used as a second reader for off-hours CTPA studies.
European Congress of Radiology, 4-8 March 2010, Vienna, Austria.

Wittenberg R, Peters JF, Weber M, Lely RJ, Cobben LPJ, Prokop M, Schaefer-Prokop CM.
Stand-alone performance of a Computer-Assisted Detection prototype for detection of acute pulmonary embolism: A multi-institutional comparison.
Radiological Society of North America, 29 November-4 December 2010, Chicago, United States of America.

Wittenberg R, Berger FH, Peters JF, Weber M, van Hoorn F, Beenen LFM, van Doorn MMAC, van Schuppen J, Zijlstra IJAJ, Prokop M, Schaefer-Prokop CM. Impact of a computer-assisted detection prototype on the detection of acute pulmonary embolism used as second reader for off-hours CTPA studies.
Radiologendagen, 17-18 September 2009, Amsterdam, The Netherlands.

Wittenberg R, Peters JF, Sonnemans JJ, Prokop M, Schaefer-Prokop CM. Computer-assisted detection of pulmonary embolism: evaluation of pulmonary CT angiographies performed in an on call setting.
2nd World Congress of Thoracic Imaging, 30 May-2 June 2009, Valencia, Spain.

Wittenberg R, Peters JF, Sonnemans JJ, Bipat S, Prokop M, Schaefer-Prokop CM. Impact of image quality on the performance of a computer assisted detection prototype for pulmonary embolism.
2nd World Congress of Thoracic Imaging, 30 May-2 June 2009, Valencia, Spain.

Conference proceedings - poster presentations

Wittenberg R, van Vliet JW, Ghaye B, Peters JF, Schaefer-Prokop CM, Coche E. Comparison of automated 4-chamber cardiac views versus axial views for measuring right ventricular enlargement in patients with suspected pulmonary embolism.
European Congress of Radiology, 4-8 March 2010, Vienna, Austria.

Wittenberg R, Berger FH, Peters JF, Weber M, van Hoorn F, Beenen LFM, van Doorn MMAC, van Schuppen J, Zijlstra IJAJ, Prokop M, Schaefer-Prokop CM. Impact of a CAD prototype on the detection of acute pulmonary embolism used as second reader for off-hours CTPA studies.
Radiological Society of North America, 29 November-4 December 2010, Chicago, United States of America.

Prices

Best abstract and presentation award

Wittenberg R, Berger FH, Peters JF, Weber M, van Hoorn F, Beenen LFM, van Doorn MMAC, van Schuppen J, Zijlstra IJAJ, Prokop M, Schaefer-Prokop CM. Impact of a computer-assisted detection prototype on the detection of acute pulmonary embolism used as second reader for off-hours CTPA studies.
Radiologendagen, 17-18 September 2009, Amsterdam, The Netherlands.

List of publications

Dankwoord

Curriculum vitae

Dankwoord

Nu ik aan het schrijven van mijn dankwoord begin, betekent het dat het einde van mijn promotie echt in zicht is. En dit is mede dankzij velen van jullie, want promoveren blijft teamwork! Daarom wil ik iedereen die bijgedragen heeft aan dit proefschrift bedanken voor hun enthousiasme en inzet.

Dr. Schaefer-Prokop, beste Cornelia, jouw liefde voor het vak en enthousiasme voor de wetenschap is bewonderenswaardig. De eerste paar maanden tijdens mijn promotie heb je je volledig ingezet om mij longembolieën te leren detecteren totdat we er allebei van konden dromen. Maar ook daarna stond jij 7 dagen in de week voor mij klaar. Niet alleen voor mijn promotie, maar ook voor mijn opleiding heb je veel voor mij gedaan waar ik je erg dankbaar voor ben. Ik hoop dan ook dat we nog vaak zullen samenwerken!

Prof. Prokop, beste Mathias, vanaf dag 1 heb jij mij opgenomen in jullie 'pulmo groep'. Hier heb ik leren presenteren en discussiëren over mijn onderwerp, waardoor ik sterker in mijn schoenen stond tijdens congressen. Ook je vernieuwende inzichten en je inzet bij het ontwikkelen van een referentie standaard voor longembolieën hebben er toe geleid dat mijn promotietraject een succes is geworden.

Prof. Laméris, bedankt voor het beoordelen van mijn manuscript en het vertrouwen dat u in mij had dat ik naast mijn opleiding ook kon promoveren.

Dr. Peters, beste Joost, wij zijn een goed voorbeeld van dat teamwork: ik de klinische kant van de studie en jij de software. En dit heb je geweten ook! Heel wat uurtjes heb jij mij aan de telefoon door de software 'heen gepraat' als ik er niet uit kwam of updates moest installeren. Ook de keren dat je in de auto bent gesprongen om mij te helpen met al die (voor mij dan) computer wartaal zal ik niet vergeten.

Prof. Van Ginneken, beste Bram, met jouw kritische blik en ervaring hield jij mij scherp tijdens de wekelijkse besprekingen. Bedankt dat je wilt plaats nemen in mijn promotiecommissie.
Ook Dr. Hartmann wil ik bedanken voor het plaats nemen in mijn promotiecommissie en de samenwerking de afgelopen jaren.

De overige leden van de beoordelingscommissie, Prof. Mali, Prof. Lammers, Prof. Viergever en Prof. Reekers, bedankt voor het beoordelen van mijn proefschrift.

Dr. Weber, dear Michael, I've been spamming your inbox with all kinds of questions and it was never too much, even if you had to explain the same thing to me again and again. With your help, I didn't feel so lost in the statistical universe anymore and I looked always forward to our next email conversation. Thank you very much for all your help and patience!

Omdat mijn onderzoek een samenwerking was tussen het AMC en UMC Utrecht heb ik het voordeel gehad in beide onderzoeksgroepen te mogen meedraaien:
Beste collega onderzoekers uit het AMC. Na even zoeken, had ik mijn plek bij jullie gevonden. Bedankt voor al jullie interesse en niet te vergeten de steun en gezelligheid tijdens congressen.
Beste collega onderzoekers van het Image Science Institute, bedankt dat jullie tijdens de wekelijkse 'pulmo meeting' en congressen altijd voor mij klaar stonden en meedachten met mijn onderzoek.
Beste collega onderzoekers uit het UMC Utrecht. Ook al was ik niet zo vaak op de afdeling aanwezig, iedere keer weer werd ik met een warm welkom ontvangen. Ik keek altijd weer uit naar het regelmatige 'taart-half uurtje' en natuurlijk alle borrels en etentjes met elkaar. Gezellig dat we nu (bijna) allemaal in opleiding zijn en elkaar nog vaak tegen zullen komen!

Veel dank gaat uit naar de volgende 'observers', die zelfs na het beoordelen van honderden longembolie scans zich niet uit het veld lieten slaan: Ludo Beenen, Ferco Berger, Inge van den Berk, Nicole Freling, Rutger Lely, Martine van Doorn, Bartjan de Hoop, Frans van Hoorn, Karin Horsthuis, Cécile Ravesloot, Joost van Schuppen en IJsbrand Zijlstra.

Beste collegae uit Amersfoort, naast mijn opleiding mijn promotie afmaken is niet altijd makkelijk geweest, maar door de prettige werksfeer, de gezellige vrijdagmiddag borrels en leuke uitjes buiten het werk, is het voor mij een leuke tijd geweest. Dank daarvoor.

Lieve Fan, Mar en Roos. Iedere keer weer zorgen jullie voor afleiding tussen al dat werken en promoveren door. Even geen gepraat over geneeskunde, maar heerlijk mezelf zijn. Iedere keer is weer een feestje en ik hoop dat er nog velen mogen volgen!

Lieve Anne-Hilde, Jana, Manon en Marieke. Tijdens alle ups and downs tijdens mijn studie, mijn werk en het promoveren waren jullie er om het samen te beleven, soms elkaar er doorheen te slepen of afleiding te zoeken. Het is leuk te zien hoe iedereen binnen de geneeskunde zijn eigen weggaat, maar altijd voor elkaar klaar blijft staan. Hopelijk volgen dan ook nog vele etentjes!

Ik kan geen betere paranimfen wensen dan jullie als vriendinnen en collega onderzoekers ineen:
Lieve Manon, 4 jaar geleden op dezelfde 2 onderzoeken gesolliciteerd en allebei aangenomen voor onze eerste keus, wie had dat gedacht! Jij in opleiding in het AMC en ik in het UMC, kunnen we mooi verhalen uitwisselen bij een goed glas wijn.
Lieve Marieke, heerlijk om samen te sporten en ons hart te luchten over het promoveren en de opleiding. Het gaat ook jou lukken om straks naast je opleiding je promotie af te ronden en als je dan straks longarts bent, kun je me altijd bellen voor een HRCT-tje!

Ik heb teveel lieve vrienden en (schoon) familie om allemaal apart op te noemen. Ook al is het soms moeilijk om uit te leggen wat ik precies gedaan heb de afgelopen 4 jaar, jullie waren altijd geïnteresseerd, een luisterend oor en vol vertrouwen dat het allemaal goed zou komen. Ik ben erg blij met zoveel steun om mij heen.

Lieve Patrick. Sinds je terug bent uit Engeland spreek ik je gelukkig vaker en heb je veel tijd en energie gestoken in het corrigeren van mijn Engelse artikelen en presentaties. Ik ben erg blij en dankbaar dat je met veel geduld en regelmaat mij de Engelse taal een beetje hebt kunnen bijbrengen!

Ook mijn ouders mogen natuurlijk niet ontbreken in het dankwoord. Op de middelbare school riepen jullie al dat onderzoek doen echt iets voor mij zou zijn. In al mijn keuzes die ik in de tussentijd heb gemaakt hebben jullie mij gesteund, zodat ik zelf mijn weg naar de radiologie heb gevonden. Ik kan altijd op jullie rekenen en dat waardeer ik enorm.

Lieve Jeroen, zonder jouw motiverende woorden en steun geen proefschrift. Ook naast mijn werk ben jij er altijd om mij te compenseren: als ik mij zorgen maak weet jij de dingen te relativeren en als ik aan het stressen ben zorg jij voor ontspanning. Jij bent er altijd voor mij en ik kijk nu al uit naar alle mooie dingen die wij samen nog gaan beleven!

Curriculum vitae

Rianne Wittenberg was born on the 1st of November 1982 in Hilversum, The Netherlands and grew up in Kortenhoef. After graduating for secondary school at the Comenius College in Hilversum in 2000, she started her medical education at the University of Amsterdam. During her study from 2002-2004 she worked at the department of Radiology at the Academic Medical Centre Amsterdam assisting the making of conventional chest X-rays on the Intensive Care. In the fourth year of her medical education she worked at this centre as a research student under supervision of Dr. R.R. van Rijn. She compared direct X-ray radiogrammetry versus dual energy X-ray absorptiometry to assess bone density in children treated for acute lymphoblastic leukaemia and growth hormone deficiency, which resulted in her first publication in Pediatric Radiology in 2006. After she obtained her medical degree in 2006 she worked at the Children's Public Health Service in Amsterdam for 1,5 year. In April 2008 she started as a PhD student at the UMC Utrecht (Prof Prokop) and the AMC Amsterdam (Prof Laméris, Dr. Schaefer-Prokop), which resulted in this thesis. While she finished this thesis, she started her radiology residence at the Meander Medical Centre (Dr. Baarslag) in April 2010, which she will continue at the UMC Utrecht (Prof. van Schaik).