

# Automatic detection of lytic and blastic thoracolumbar spine metastases on computed tomography

Matthias Hammon · Peter Dankerl · Alexey Tsymbal ·  
Michael Wels · Michael Kelm · Matthias May ·  
Michael Suehling · Michael Uder · Alexander Cavallaro

Received: 12 October 2012 / Revised: 6 December 2012 / Accepted: 19 December 2012 / Published online: 9 February 2013  
© The Author(s) 2013. This article is published with open access at Springerlink.com

## Abstract

**Objective** To evaluate a computer-aided detection (CADe) system for lytic and blastic spinal metastases on computed tomography (CT).

**Methods** We retrospectively evaluated the CADe system on 20 consecutive patients with 42 lytic and on 30 consecutive patients with 172 blastic metastases. The CADe system was trained using CT images of 114 subjects with 102 lytic and 308 blastic spinal metastases. Lesions were annotated by experienced radiologists. Detected benign lesions were considered false-positive findings. Detector sensitivity and the number of false-positive findings were calculated as the criteria for detector performance, and free-response receiver operating characteristic (FROC) analysis was conducted. Detailed analysis of false-positive and false-negative findings was performed.

**Results** Algorithm runtime is  $3 \pm 0.5$  min per patient. The system achieves a sensitivity of 83 % at 3.5 false positives per patient on average for blastic metastases and a sensitivity of 88 % at 3.7 false positives for lytic metastases. False positives appeared predominantly in the area of degenerative changes in the case of the blastic metastasis detector and in osteoporotic areas in the case of the lytic metastasis detector.

**Conclusion** The CADe system reliably detects thoracolumbar spine metastases in real time. An additional study is planned to evaluate how the bone lesion CADe system

improves radiologists' accuracy and efficiency in a clinical setting.

## Key Points

- Computer-aided detection (CADe) of bone metastases has been developed for spinal CT.
- The CADe system exhibits high sensitivity with a tolerable false-positive rate.
- Analysis of false-positive detection may further improve the system.
- CADe may reduce the number of missed spinal metastases at CT interpretation.

**Keywords** Computer-aided detection (CADe) · Bone metastases · Computed tomography (CT) · Spine · Clinical decision support

## Abbreviations

CADe	computer-aided detection
CT	computed tomography
FDG	fluorodeoxyglucose
FN	false negative
FP	false positive
FROC	free-response receiver operating characteristic
MRI	magnetic resonance imaging
PET-CT	positron emission tomography–computed tomography
TN	true negative
TP	true positive

The study was presented at ECR 2012 (abstract B-0603).

M. Hammon (✉) · P. Dankerl · M. May · M. Uder · A. Cavallaro  
Department of Radiology, University Hospital Erlangen,  
Maximiliansplatz 1,  
91054 Erlangen, Germany  
e-mail: matthias.hammon@uk-erlangen.de

A. Tsymbal · M. Wels · M. Kelm · M. Suehling  
Corporate Technology Division, Siemens AG, San Carlos Street 7,  
91058 Erlangen, Germany

## Introduction

Various types of malignant primary tumours spread to the skeletal system. Most patients die not because of the growth of the primary cancer, but because of its spread to other sites [1]. Bone metastases portends a poor survival with a median

of less than 6 months [2] and, therefore, are of critical importance for oncological patients. Furthermore, metastasis to the skeletal system is a frequent and serious complication of various cancer types with high incidence and prevalence, such as breast cancer, colon cancer and lung cancer [3]. Accurate detection of bone metastases is thus an important task in a radiologist's daily routine, because it provides valuable clinical information and enables the timely choice of systemic or local therapy such as surgical intervention or radiation. There are different manifestations of bone metastases; they appear lytic, blastic (sclerotic) or show a mixed appearance [4]. In particular, lytic metastases within the vertebral column can cause pathological fractures, severe pain and spinal cord compression, potentially together with neurological impairment [1]. Moreover, the presence of bone metastasis can be an important prognostic factor of whether the patient will benefit from chemotherapy which is often associated with impairing side effects [5].

Nowadays there are many radiological methods available for examination of the skeletal system for osseous metastases, such as computed tomography (CT), magnetic resonance imaging (MRI), skeletal scintigraphy or positron emission tomography–CT (PET-CT). However, in clinical routine, the initial staging and especially the follow-up examinations of oncological patients often include CT imaging only. CT demonstrates superior bony detail, allowing early detection of bone metastases [6]. Nevertheless, it is challenging and time consuming to detect bone lesions at an early stage on CT images especially when a variety of benign osseous lesions with a lytic or sclerotic appearance, such as osteoporosis or degenerative changes, are present. Moreover, it has been postulated that skeletal metastases are at risk of being missed because bone windows are underutilised in a radiologist's daily routine [7].

Therefore, reliable and reproducible automatic detection of spinal metastases in CT images can be regarded as a desperately needed and useful tool in the diagnosis, staging and treatment monitoring of cancer patients. It could assist the reader in the final decision making as it indicates suspicious osseous regions for further consideration. Furthermore, it can be part of an often demanded multipurpose computer-aided detection system [8]. To efficiently utilise the software in the daily routine, it is crucial to achieve satisfactory sensitivity. On the other hand, the number of false-positive results has to be minimised, and the system performance should be as fast as possible in order not to markedly delay or prolong the reporting.

In this study, the performance of a piece of automatic detection software for lytic and blastic spinal metastases on CT images was evaluated. The software could potentially assist radiologists in detecting thoracolumbar spine metastases.

## Materials and methods

The institutional review board of the University of Erlangen-Nuremberg approved this study and waived the need for informed consent.

The software was developed within the framework of the German Theseus-Medico research programme, a recently completed 5-year nationwide multi-centre research project [9, 10]. Physicians, university health-care professionals and computer scientists collaborated in this joint venture. The Theseus-Medico software platform was developed in the course of this programme to support physicians in accurate and efficient patient diagnostics and patient monitoring in various application areas including spinal metastases detection.

### Patient population

For this retrospective study the Radiological Information System (RIS) was used to search for patients with reported lytic and/or blastic bone metastases in the thoracolumbar spine who underwent CT imaging from 1 January 2011 to 31 December 2011. A radiological resident with 2 years' experience and a board-certified radiologist with 15 years' experience collected and annotated the retrieved CT data in consensus. Consecutive patients with at least one confirmed malignant lytic lesion larger than  $0.5\text{ cm}^3$  (equal to a diameter of roughly 10 mm) or one malignant blastic lesion larger than  $0.3\text{ cm}^3$  (equal to a diameter of roughly 8 mm; similar to O'Connor et al. [11] and Wiese et al. [12]) in the vertebral bodies of the thoracolumbar spine were selected. Instead of completely excluding patients with metal implants or traumatic fractures from the experimental analysis, only specific vertebral bodies displaying metal implants (e.g. screws;  $n$  (vertebral bodies)=4), compression fractures ( $n=7$ ) or kyphoplasty material ( $n=5$ ) were excluded. The first 50 subjects who matched these parameters were included (31 women, 19 men; mean age, 58 years; range, 32–87 years). Twenty of these patients showed lytic lesions; 30 patients showed blastic lesions. The underlying primary cancer was breast cancer ( $n=11$ ), plasmacytoma ( $n=8$ ), prostate cancer ( $n=5$ ), melanoma ( $n=5$ ), renal cell carcinoma ( $n=4$ ), pancreatic carcinoma ( $n=3$ ), lung cancer ( $n=3$ ), lymphoma ( $n=2$ ), colorectal carcinoma ( $n=2$ ), oesophageal carcinoma ( $n=2$ ) and other types of cancer ( $n=5$ ). All patients underwent contrast-enhanced thoraco/abdominal CT imaging.

Lesion annotation was performed by positioning a 3D rectangular bounding box around each lesion and assigning the lesion's type. Lesion type was verified by all available information concerning the patient such as patient history, MRI, skeletal scintigraphy, PET-CT and histology. Although histological confirmation was not available for all cases, the

lesion annotations were verified in the clinical routine until an adequate assessment was possible.

### CT technique

CT was performed with a Somatom Sensation® 64-detector row system (Siemens AG, Erlangen, Germany) with the following parameters: craniocaudal thoraco/abdominal CT data acquisition, 120 kV, Care Dose® (Siemens AG, Erlangen, Germany); pitch, 0.9; collimation, 0.6 mm; section thickness, 3 mm; hard recon kernel. Images were acquired at the portal-venous contrast agent phases (intravenous application of weight-adapted, warmed Imeron® 400 (Bracco Imaging, Konstanz, Germany) followed by a saline flush with a flow rate of 3 mL/s through a 20-gauge catheter in an antecubital vein.

### Computer-aided detection system

The detection process starts with the automatic detection of vertebral bodies (Fig. 1). In CT data, vertebral bodies can be reliably detected, for example, using iterated marginal space learning [13]. Vertebral bodies are highlighted on the CT image and restrict the search space for subsequent lesion detection. The images of vertebral bodies are spatially normalised to have the same orientation and extension and are inputted into a cascade detector consisting of *three* random forest-based discriminative models [14] each working on a selection of features describing the individual lesion centre candidates. The first model in the detector exploits a limited set of features which includes low-level 3D Haar-like features. The other subsequent two models exploit the full heterogeneous set of features, which are of different nature and describe various characteristics of the suspected lesion [13, 15, 16]. The lesion centre detector cascade provides rough to coarse lesion detection, which starts with a large set of suspicious lesion-like structures and ends with a reduced set of likely malignant clinically important findings. After having detected the centres of likely bone metastases, the system uses a patient-specific estimation of the spongiosa's

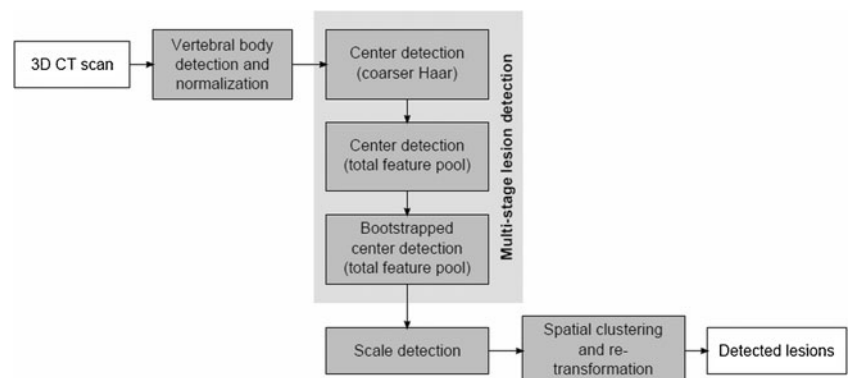
intensity distribution within all segmented vertebral bodies to additionally reject lesion candidates whose centre voxels do not sufficiently differ from the surrounding spongy bone tissue. The remaining findings are refined with scale estimates by an additional scale detector, grouped together with hierarchical agglomerative clustering, transformed back to the coordinate system of the original CT data and ultimately displayed to the clinician as bounding boxes on the CT image.

In our study, the detection cascade, both for lytic and blastic metastasis detection, was parameterized so as to have, per each vertebral body, up to 1,000 findings at the output of the first model, up to 200 candidates after applying the second model, and up to 100 voxel candidates at the cascade output and input for agglomerative clustering, both in detection training and its application. This setting has been proven to be the best, confirmed with cross validation on the training data. Owing to the probabilistic nature of the models, the system can be tuned to operate at different levels of sensitivity and specificity. In our study, the operating point (likelihood threshold) was chosen so that the expected amount of false positives per patient did not exceed 4, which was considered to be a tolerable amount of false positives in the clinical routine.

The detection of lytic and blastic metastases is independent from each other and can be initiated by a clinician on demand. The detection processes in the search for lytic and blastic metastases are identical and correspond to that presented in Fig. 1. The two respective detector training processes differ only in the specific sets of positive and negative bone lesions provided to train each component of the detector. Training of such a detector system was discussed in more detail by Wels et al. [17], with an application in lytic metastases detection.

An important peculiarity in the detector cascade is the classification model, which is trained to explicitly differentiate between the clinically interesting malignant findings and similar looking benign lesions (e.g. osteophytes for blastic lesions and osteoporotic areas or the basivertebral vein for lytic ones). The set of negative candidates to train

**Fig. 1** Machine learning-based framework for the automatic detection of lytic and blastic spinal metastases in CT images. The flow chart shows the different steps of the system. 3D three-dimensional, CT computed tomography



the model is sampled so as to include a certain share of benign degenerations, which represent typical false positives. This solution was recently implemented as an extension of Wels et al.'s approach [17] and was proven to reduce the number of clinically uninteresting, false-positive, findings.

The CADe system was trained on lesions with an ellipsoid volume larger than  $0.3 \text{ cm}^3$  for blastic metastases (similar to Wiese et al. [12]) and  $0.5 \text{ cm}^3$  for lytic metastases (similar to O'Connor et al. [11]). Training data were collected by searching the RIS for patients with reported lytic and/or blastic bone metastases in the thoracolumbar spine who underwent CT from 1 January 2009 to 31 December 2010. Thoraco/abdominal CT images of 114 subjects (67 women, 47 men) were annotated for training of the detectors. A group of 41 patients showed 102 lytic metastases and 73 patients showed 308 blastic metastases, respectively. Benign bone lesions such as osteophytes ( $n=576$ ), degenerative sclerosis ( $n=367$ ), Schmorl's nodules ( $n=146$ ), osteoporotic areas ( $n=96$ ), haemangiomas ( $n=21$ ) as well as the basivertebral vein were annotated to train the last classification model in the detector cascade. A standard personal computer (Dual Core Xeon 2.66 GHz; Windows XP, 32 bit) was used for evaluation of the system.

#### Statistical analysis

Software for the performance evaluation and statistical analysis of lesion detection was implemented as a part of the CADe framework. The system is able to automatically generate a performance report for a set of CT images with annotated lytic and blastic metastases. Sensitivities per lesion and per patient, positive predictive values per lesion and per patient, the number of false positives per patient and the mean runtime (with standard deviation) were calculated for the two detectors (lytic and blastic) to assess the predictive performance of the CADe system. Free-response receiver operating characteristic (FROC) curves were also generated.

A lesion is considered to be detected as soon as a detection result's centre lies within the bounding box of the expert-annotated bone lesion. Further, several detections within a single vertebral body that is completely affected by either lytic or blastic metastases are not counted as false-positive detections. Missed detections of lytic or blastic metastases whose ellipsoidal volume is smaller than  $0.5$  and  $0.3 \text{ cm}^3$  respectively (which correspond to the thresholds in previously published related studies [11, 12]) are not counted as false-negative detections. Such annotations were excluded from evaluation and do not affect the counts of false positives (FP), true positives (TP), false negatives (FN) and true positives (TP).

## Results

Sample lytic and blastic metastases detections of our CADe system are shown in Fig. 2. Figure 3 includes the curves of the FROC analysis of the two detectors (lytic and blastic metastases), on the test patients. As one see, both the FROC curves plateau at a certain level of sensitivity, around 90 %. This occurs due to the fact that a multi-level model involving a detector cascade is used. The classification model at each cascade level has its own sensitivity (which is high enough but is less than 100 %), which lowers the overall reachable sensitivity of the system, but makes the amount of false positives clinically tolerable at the operating points of interest, comparing with single-level solutions.

#### Detection of blastic spinal metastases

The average runtime of the CADe system is  $95 \pm 12$  s per patient. On the 30 test cases the system achieves a per-lesion sensitivity of 83 % and a per-patient sensitivity of 80 % at 3.5 false-positive detections per patient. Thus, 143 of the 172 annotated metastases were successfully detected. The number of false-positive detections per patient varies from 0 to 26. The per-lesion and per-patient positive predictive values are 58 % and 65 % (see Table 1 for a complete summary of the results).

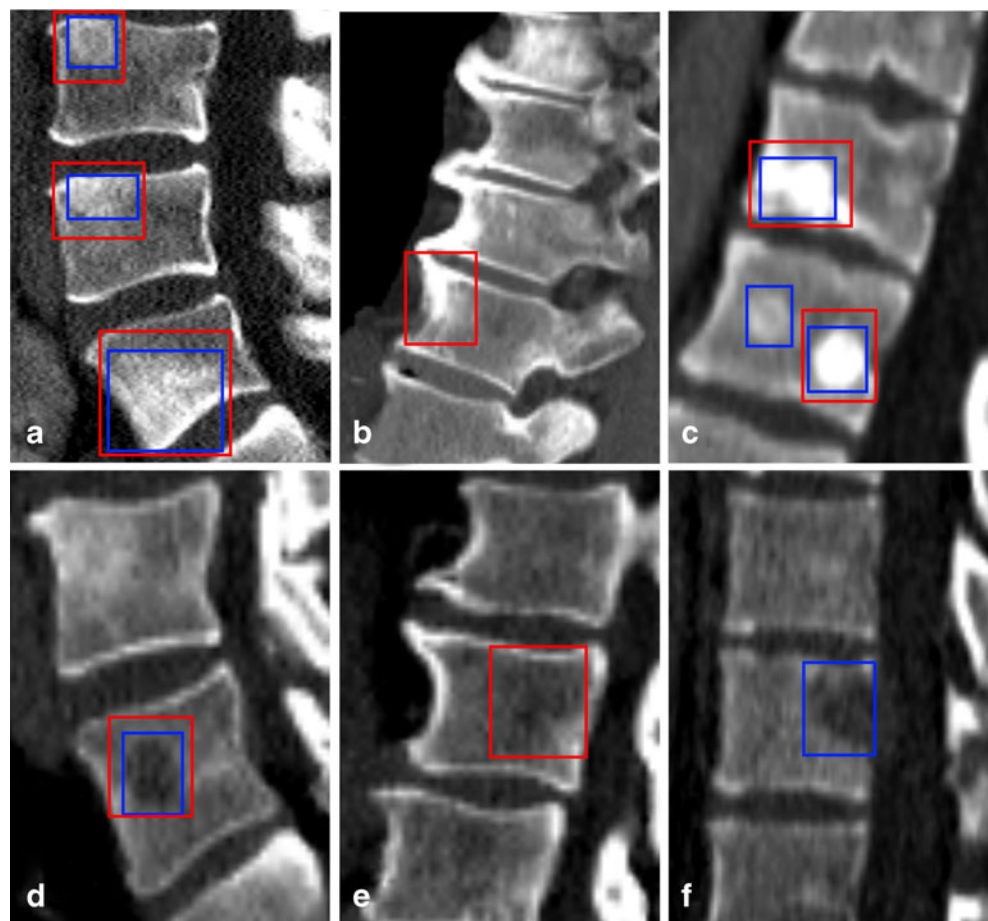
The majority of the 102 false-positive detections were caused by degenerative changes (see Table 2 for detailed information). Figure 4 shows examples of false-positive detections. All false-positive detections were located inside the vertebral bodies, mostly on or near the surface/end plates (Table 2). Of the 29 false-negative detections, 24 misdetections had a maximum diameter not larger than 1 cm, indicating the trend that smaller blastic metastases are more likely to be missed. Figure 5 shows examples of false-negative detections.

#### Detection of lytic spinal metastases

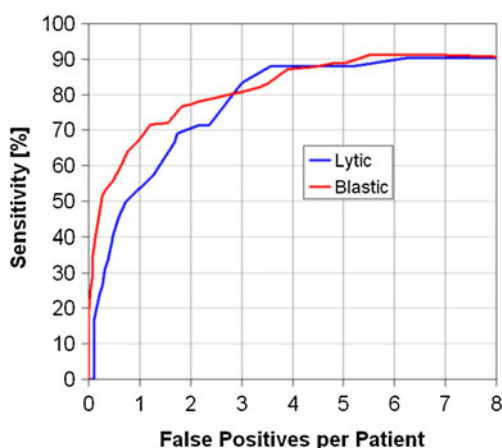
The average runtime of the CADe system is  $87 \pm 18$  s per patient. On the 20 test cases the system achieves a per-lesion sensitivity of 88 % and a per-patient sensitivity of 93 % at 3.7 false positives per patient. Thus, 37 of the 42 annotated metastases were successfully detected. The number of false-positive detections per patient varies from 0 to 11. The per-lesion and per-patient positive predictive values are 35 % and 49 %, respectively (see Table 1 for a complete summary of the results).

The majority of the 70 false-positive detections were caused by osteoporotic changes (see Table 2 for detailed information). Figure 4 shows examples of false-positive

**Fig. 2** Examples of detection results on CT images (sagittal plane). True-positive (a), false-positive (b) and false-negative (c) detections of blastic metastases in the thoracolumbar spine are shown in the upper row; the lower row shows analogous detections of lytic metastases (d–f). The blue boxes represent the ground-truth annotations; the red boxes represent the detection results of the CAde system



detections. A set of 7 (10 %) false-positive detections were located outside the vertebra (Table 2). Contrary to blastic metastasis, four of the five false-negative detections had a maximum diameter larger than 1 cm. Figure 5 shows examples of false-negative detections.



**Fig. 3** Free-response receiver operating characteristic (FROC) curves showing the per-lesion sensitivity and the number of false-positive detections per patient of the CAde system for lytic (blue line) and blastic (red line) metastases in the vertebral bodies of the thoracolumbar spine

## Discussion

As demonstrated by this study, the computer-aided detection (CAde) system quickly detects bone metastases in the thoracolumbar spine with a sensitivity of 88 % for lytic metastases and 83 % for blastic ones. The number of false-positive detections was slightly higher in the case of the detection of lytic metastases (3.7 versus 3.5 per patient).

Influenced by the persistent trend of an ever-increasing image data volume and the associated rising workload of radiologists [18, 19], there is a mounting call for computer-based support that will assist the reading process [20]. Various CAde systems for multiple clinical tasks have already been developed in an attempt to address this challenge. Some of them are already successfully used in daily clinical routine, with others in the process of gaining clinical acceptance. CAde systems for lung, breast, colon, liver and prostate cancer as well as for coronary stenosis and pulmonary embolism have been studied recently and found clinical acceptance [21–27].

The sensitivity achieved and the number of false-positive detections of the system under evaluation are promising. Therefore we propose to further refine the system and to evaluate it in a clinical setting, so that it could become a

**Table 1** Results of the automatic detection of blastic and lytic spinal metastases on CT images

	Blastic metastasis detector	Lytic metastasis detector
Total volumes	30	20
Total annotations	172	42
Total algorithm detections	245	107
Total true positives	143	37
Total false positives	102	70
Total false negatives	29	5
False positives per patient	3.5	3.7
Sensitivity per lesion (%)	83	88
Sensitivity per patient (%)	80	93
Positive predictive value per lesion (%)	58	35
Positive predictive value per patient (%)	65	49
Mean runtime $\pm$ SD (s)	93 $\pm$ 12	87 $\pm$ 18

component of an often-demanded multipurpose CADe [8] which is expected to improve radiologists' accuracy and productivity [28]. For the purpose of displaying clinical feasibility and seamless integration into the reading process, as recommended by van Ginneken et al. [29] we examined the software in a cohort of consecutive patients with different types of primary cancers. As analysis of the results suggests, the major application of the software could be in guiding the radiologist to suspicious areas for further consideration, as early-stage bone metastases in particular are difficult to detect on CT images even for experienced radiologists [7].

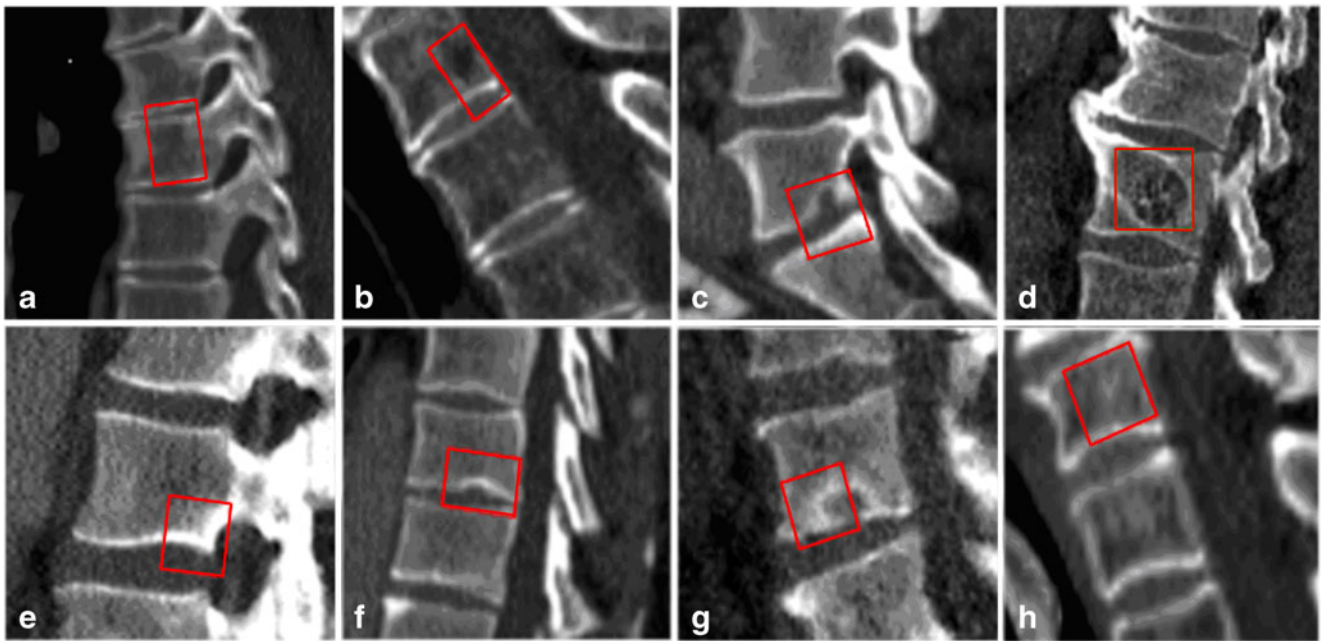
O'Connor et al. [11] previously described a CADe system for lytic bone metastases in the thoracolumbar spine. They observed that 27 % of false-positive results were located in the region of the intervertebral disc. Similar observations were made in earlier editions of our system. In order to reduce the number of false-positive detections, we applied a segmentation of the vertebral bodies in a preprocessing step to exclude the intrinsically low-attenuating intervertebral discs. O'Connor et al. [11] used 50 cases (including 28 lesions) with a section thickness of 5 mm, leading to volume averaging, which caused the segmentation of undesired structures. We

used images with a 3-mm section thickness sagittal reconstruction of routine thoraco-abdominal CT images. O'Connor et al. [11] mentioned false-positive findings caused by non-pathological, low-attenuating structures, such as the basivertebral vein and by detections outside the vertebra. We met this challenge by limiting the search space to the vertebral body because it can be robustly segmented and because most metastases are located in this part of the vertebra [30]. Additionally, the automatic segmentation of the vertebral bodies allows searching in the region of the basivertebral vein, which is known to always reside in the same segment of the vertebral body, to be avoided. Owing to varying conditions, the two systems are difficult to compare. Our system generated fewer false-positive lesions (3.7 versus 4.5 per patient) whereas the sensitivity (88 %) was located between their training (83 %) and test set sensitivity (94 %). In comparison to O'Connor et al.'s work [11], the software was applied to all consecutive CT examinations and only single vertebral bodies showing hardware, such as screws, compression fractures or kyphoplasty material, were excluded from evaluation. No subject was excluded because of bad performance, unsatisfactory image quality, extensive disease or extensive degeneration.

Wiese et al. [12] described a CADe system for blastic bone metastases in the spine. In comparison to this study we were able to achieve a significantly better sensitivity at a smaller number of false-positive detections when concentrating on the vertebral body. Wiese et al. [12] did not analyse the false-positive detections, although according to our experience, degenerative changes such as osteophytes are mostly responsible for the high numbers of false-positive detections. We exploited a larger set of blastic metastases and, as described before, our system was trained to explicitly avoid detection of non-malignant lesions to reduce the false-positive rate. This has proven to be an efficient solution, eliminating obviously benign findings and significantly reducing the number of false positives, on average by up to two per patient, both for blastic and lytic metastases. This solution can be potentially useful for other similar CADe systems where the likely false positives include a separate group with a specific location and/or appearance and can be annotated and incorporated in the learning process in order to improve the predictive performance of the system.

**Table 2** Analysis of the false-positive results of the spinal metastases detection software

	False-positive results	
	Blastic metastasis detector ( $n=102$ )	Lytic metastasis detector ( $n=70$ )
Location	All inside the vertebra, predominantly close to the end plates	Predominantly in the vertebra ( $n=63$ ; 90 %), some outside ( $n=7$ ; 10 %)
Type	Mainly osteophytes ( $n=42$ ; 41 %) and degenerative sclerosis ( $n=39$ ; 38 %), some Schmorl's nodules ( $n=5$ ; 5 %) and non-classifiable image artefacts, respectively ( $n=16$ ; 16 %)	Mainly osteoporotic areas ( $n=54$ ; 77 %), some basivertebral veins ( $n=5$ ; 7 %), haemangiomas ( $n=3$ ; 4 %) and a Schmorl's nodule ( $n=1$ ; 1 %)



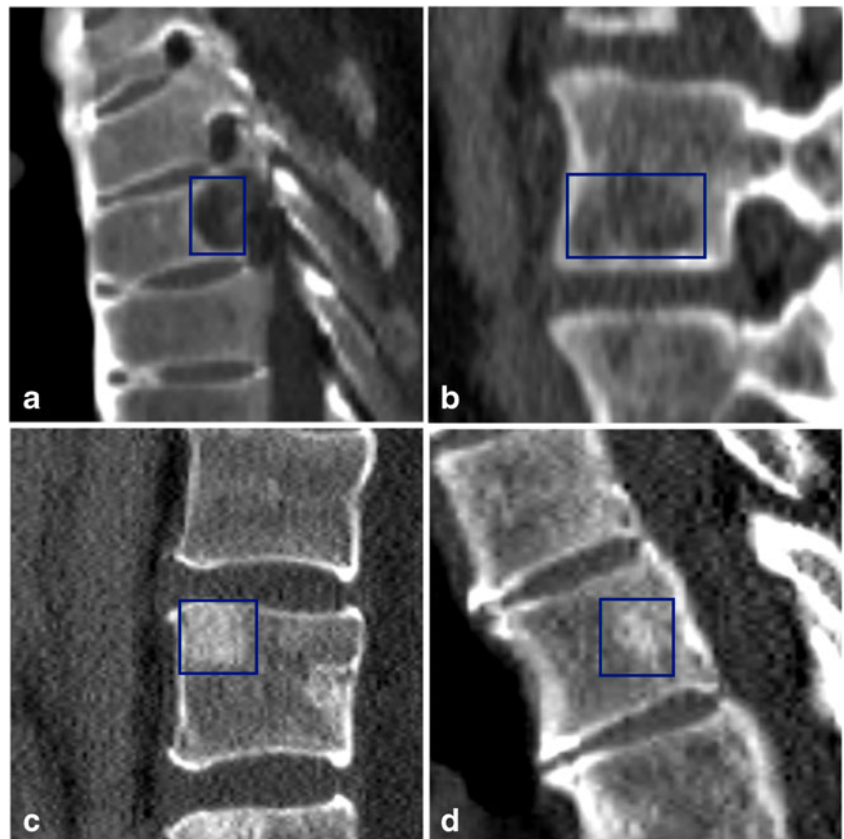
**Fig. 4** Examples of false-positive detections (sagittal plane). False-positive detections of the lytic metastases detector are shown in the *upper row* (a an osteoporotic area, b a basivertebral vein, c a Schmorl's nodule, d a haemangioma); the *lower row* shows analogous false-

positive detections of the blastic metastases detector (e an osteophyte, f a degenerative sclerosis, g a Schmorl's nodule, h a non-classifiable). The *red boxes* represent the detection results of the CADE system

The runtimes for the software introduced by O'Connor et al. [11] and Wiese et al. [12] are unknown. As the overall

runtime of our system is 3 min per patient it does not delay the reading process which is important for clinical

**Fig. 5** Examples of false-negative detections (sagittal plane). False-negative detections of the lytic metastases detector are shown in the *upper row* (a a lytic metastasis close to the basivertebral vein, b a lytic metastasis falsely detected as an osteoporotic area); the *lower row* shows analogous false-negative detections of the blastic metastases detector (c a blastic metastasis close to the end plate, d a subtle blastic metastasis). The *blue boxes* represent the ground-truth annotations



feasibility. Usually, the CADe system would be able to generate its detections while the radiologist is reading the soft tissue window. Because the system is completely automatic, it can also be launched in an offline mode, before the case is opened for reading.

Although alternative methods for the detection of bone metastases such as bone scintigraphy, PET-CT and MRI have often been proven to provide better sensitivity, we are studying CT because in daily routine a radiologist must also detect bone metastases on the more commonly acquired CT images. Machine learning-based CADe systems could assist the reader in recognising suspicious osseous areas on CT. Subsequently, in the case of high-risk patients, more sensitive or specific methods could be provided to detect and treat osseous metastases at an early stage. Interestingly, it was observed that PET and bone scintigraphy also have limitations concerning the detection and the assessment of bone metastases. Thus, some benign bone lesions show a high accumulation of fluorodeoxyglucose (FDG) [31] and blastic metastases are often not detectable by PET because they show low metabolic activity [32]. In the case of usual technetium 99-m ( $^{99m}\text{Tc}$ )-based bone scintigraphy, lytic bone lesions related to multiple myeloma do not show uptake of the radioisotope [33].

The FROC curves plot the per-lesion sensitivity of each detector versus the expected number of false-positive detections per patient (Fig. 3). The system allows easy tuning to work at any operating point on these FROC curves. This is important, as long as, under some circumstances, a radiologist would find lower sensitivity with fewer false positives acceptable (e.g. in patients who show advanced degenerative changes of the spine but are not expected to show osseous metastases), whereas in some other cases exploring every single suspicious region at the expense of observing more false positives might be desirable, in order to increase the sensitivity of the reading (e.g. in an oncology patient). For images that are not likely to show many false positives (e.g. of a young patient with a spine not showing degeneration), the radiologist may demand high sensitivity along with a tolerable false-positive rate. With our system, the operating point can be changed in real time, increasing or reducing the set of detections, through thresholding their likelihood.

Our study faces some limitations that suggest directions for future work. Contrary to the alternative approaches in which the whole vertebra is considered [11, 12], the system was trained to detect lesions in the vertebral body only. We intentionally concentrated on the vertebral body because it reduces runtime and improves predictive performance. Moreover, it is known that most metastases are located in this part of the vertebra [30].

Furthermore, the system was trained and tested with lytic metastases larger than  $0.5\text{ cm}^3$  and blastic metastases larger

than  $0.3\text{ cm}^3$  which are the same threshold values as those in previously published related studies [11, 12]. To achieve satisfying results in detecting smaller lesions as well, larger patient populations with a reliable annotation of lesions of this kind are required for the training (and evaluation) of the detectors.

In the present study we concentrated on the detection of lytic and blastic metastases. We noticed that in case of a mixed metastasis, showing a lytic and a blastic portion, both detectors usually generate findings of the corresponding portion. Revision of collocated findings of this kind is a subject of our ongoing work.

It has been reported recently that CADe systems have the potential to detect up to 50 % of the lesions overlooked by human readers in the case of breast and lung cancer [34, 35]. A study evaluating how the CADe system may help in decreasing the number of missed spinal bone metastases similar to Nishikawa et al. [34] and White et al. [35] can be conducted when a big enough set of lesions that tend to be overlooked by a radiologist is collected. This is an interesting and important direction for future work.

In conclusion, the CADe system under evaluation reliably and quickly detects thoracolumbar spine metastases in CT images. It can be applied as a fully automatic preprocessing step to indicate suspicious osseous areas and thus support the radiologist. This is extremely valuable during the reading of CT images as bone metastases potentially cause severe impairment of the patient and risk being missed even by experienced radiologists. An additional study is planned to evaluate how the bone lesion CADe system improves radiologists' accuracy and efficiency in a clinical setting.

**Acknowledgments** This work was done within the framework of the Theseus-Medico research programme supported by the German Federal Ministry of Economics and Technology (project number 01MQ07026).

**Conflict of interest** A. T., M. W., M. K. and M. S. are employed by Siemens AG and were research partners during the government funded programme.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

1. Mundy GR (2002) Metastasis to bone: causes, consequences and therapeutic opportunities. *Nat Rev Cancer* 2:584–593
2. Theriault RL, Theriault RL (2012) Biology of bone metastases. *Cancer Control* 19:92–101
3. Roodman GD (2004) Mechanisms of bone metastasis. *N Engl J Med* 15:1655–1664
4. Choi J, Raghavan M (2012) Diagnostic imaging and image-guided therapy of skeletal metastases. *Cancer Control* 19:102–112



5. Kodaira M, Takahashi S, Yamada S et al (2010) Bone metastasis and poor performance status are prognostic factors for survival of carcinoma of unknown primary site in patients treated with systematic chemotherapy. *Ann Oncol* 21:1163–1167
6. Kalogeropoulou C, Karachaliou A, Zampakis P (2009) Radiologic evaluation of skeletal metastases: role of plain radiographs and computed tomography. *Bone Metastases Cancer Metastasis Biol Treat* 12:119–136
7. Hamaoka T, Madewell JE, Podoloff DA, Hortobagyi GN, Ueno NT (2004) Bone imaging in metastatic breast cancer. *J Clin Oncol* 15:2942–2953
8. Krupinski EA (2004) Computer-aided detection in clinical environment: benefits and challenges for radiologists. *Radiology* 231:7–9
9. MEDICO (2011) Intelligent searches of medical databases. Federal Ministry of Economics and Technology, Germany. <http://www.theseus-programm.de/en/920.php>. Accessed 4 Dec 2012
10. Krueger-Brand HE (2009) Web 3.0 im Dienst der Medizin. *Dtsch Aerzteblatt* 106:458–459 (in German)
11. O'Connor SD, Yao J, Summers RM (2007) Lytic metastases in thoracolumbar spine: computer-aided detection at CT-preliminary study. *Radiology* 242:811–816
12. Wiese T, Yao J, Burns JE, Summers RM (2012) Detection of sclerotic bone metastases in the spine using watershed algorithm and graph cut. *SPIE Med Imaging Proc* 8315
13. Kelm BM, Zhou SK, Suehling M, Zheng Y, Wels M, Comaniciu D (2011) Detection of 3D spinal geometry using iterated marginal space learning. *Springer Lect Notes Comput Sci* 6533:96–105
14. Breiman L (2001) Random forests. *Mach Learn* 45:5–32
15. Barbu A, Suehling M, Xu X, Liu D, Zhou SK, Comaniciu D (2010) Automatic detection and segmentation of axillary lymph nodes. *Springer Lect Notes Comput Sci* 6361:28–36
16. Wels M, Zheng Y, Carneiro G, Huber M, Hornegger J, Comaniciu D (2009) Fast and robust 3-D MRI brain structure segmentation. *Springer Lect Notes Comput Sci* 5762:575–583
17. Wels M, Kelm BM, Tsybal A et al (2012) Multi-stage osteolytic spinal bone lesion detection from CT data with internal sensitivity control. *SPIE Med Imaging Proc* 8315
18. Andriole KP, Morin RL, Arenson RL et al (2004) Addressing the coming radiology crisis—the Society for Computer Applications in Radiology transforming the radiological interpretation process (TRIP) initiative. *J Digit Imaging* 17:235–243
19. Andriole KP, Wolfe JM, Khorasani R et al (2011) Optimizing analysis, visualization, and navigation of large image data sets: one 5000-section CT scan can ruin your whole day. *Radiology* 259:346–362
20. Summers RM (2003) Road maps for advancement of radiologic computer-aided detection in the 21st century. *Radiology* 229:11–13
21. Roos JE, Paik D, Olsen D et al (2010) Computer-aided detection (CAD) of lung nodules in CT scans: radiologist performance and reading time with incremental CAD assistance. *Eur Radiol* 20:549–557
22. Blackmon KN, Florin C, Bogoni L et al (2011) Computer-aided detection of pulmonary embolism at CT pulmonary angiography: can it improve performance of inexperienced readers? *Eur Radiol* 21:1214–1223
23. Dorrius MD, Jansen-van der Weide MC, van Ooijen PM, Pijnappel RM, Oudkerk M (2011) Computer-aided detection in breast MRI: a systematic review and meta-analysis. *Eur Radiol* 21:1600–1608
24. Neri E, Faggioni L, Regge D et al (2011) CT Colonography: role of a second reader CAD paradigm in the initial training of radiologists. *Eur J Radiol* 80:303–309
25. Napel SA, Beaulieu CF, Rodriguez C et al (2010) Automated retrieval of CT images of liver lesions on the basis of image similarity: method and preliminary results. *Radiology* 256:243–252
26. Sung YS, Kwon HJ, Park BW et al (2011) Prostate cancer detection on dynamic contrast-enhanced MRI: computer-aided diagnosis versus single perfusion parameter maps. *AJR Am J Roentgenol* 197:1122–1129
27. Arnoldi E, Gebregziabher M, Schoepf UJ et al (2010) Automated computer-aided stenosis detection at coronary CT angiography: initial experience. *Eur Radiol* 20:1160–1167
28. Kotter E, Langer M (2011) Computer aided detection and diagnosis in radiology. *Eur Radiol* 21:590–592
29. Van Ginneken B, Schaefer-Prokop CM, Prokop M (2011) Computer-aided diagnosis: how to move from the laboratory to the clinic. *Radiology* 261:719–732
30. Algra PR, Heimans JJ, Valk J, Nauta JJ, Lachniet M, Van Kooten B (1992) Do metastases in vertebrae begin in the body or the pedicles? Imaging study in 45 patients. *AJR Am J Roentgenol* 158:1275–1279
31. Aoki J, Watanabe H, Shinozaki T et al (2001) FDG PET of primary benign and malignant bone tumors: standardized uptake value in 52 lesions. *Radiology* 219:774–777
32. Cook GJ, Houston S, Rubens R, Maisey MN, Fogelman I (1998) Detection of bone metastases in breast cancer by 18FDG PET: differing metabolic activity in osteoblastic and osteolytic lesions. *J Clin Oncol* 16:3375–3379
33. Angtuaco EJ, Fassas AB, Walker R, Sethi R, Barlogie B (2004) Multiple myeloma: clinical review and diagnostic imaging. *Radiology* 231:11–23
34. Nishikawa RM, Schmidt RA, Linver MN, Edwards AV, Papaioannou J, Stull MA (2012) Clinically missed cancer: how effectively can radiologists use computer-aided detection? *AJR Am J Roentgenol* 198:708–716
35. White CS, Flukinger T, Jeudy J, Chen JJ (2009) Use of a computer-aided detection system to detect missed lung cancer at chest radiography. *Radiology* 252:273–281