European Radiology (2019) 29:144–152
https://doi.org/10.1007/s00330-018-5528-6

COMPUTER APPLICATIONS

A cloud-based computer-aided detection system improves identification of lung nodules on computed tomography scans of patients with extra-thoracic malignancies

Lorenzo Vassallo[1,2] · Alberto Traverso[3,4] · Michelangelo Agnello[3] · Christian Bracco[5] · Delia Campanella[1] · Gabriele Chiara[1] · Maria Evelina Fantacci[6] · Ernesto Lopez Torres[7] · Antonio Manca[1] · Marco Saletta[7] · Valentina Giannini[1,2] · Simone Mazzetti[1,2] · Michele Stasi[5] · Piergiorgio Cerello[7] · Daniele Regge[1,2]

Received: 21 March 2018 / Revised: 27 April 2018 / Accepted: 7 May 2018 / Published online: 15 June 2018

Abstract

Objectives To compare unassisted and CAD-assisted detection and time efficiency of radiologists in reporting lung nodules on CT scans taken from patients with extra-thoracic malignancies using a Cloud-based system.

Materials and methods Three radiologists searched for pulmonary nodules in patients with extra-thoracic malignancy who underwent CT (slice thickness/spacing 2 mm/1.7 mm) between September 2015 and March 2016. All nodules detected by unassisted reading were measured and coordinates were uploaded on a cloud-based system. CAD marks were then reviewed by the same readers using the cloud-based interface. To establish the reference standard all nodules ≥ 3 mm detected by at least one radiologist were validated by two additional experienced radiologists in consensus. Reader detection rate and reporting time with and without CAD were compared. The study was approved by the local ethics committee. All patients signed written informed consent.

Results The series included 225 patients (age range 21–90 years, mean 62 years), including 75 patients having at least one nodule, for a total of 215 nodules. Stand-alone CAD sensitivity for lesions ≥ 3 mm was 85% (183/215, 95% CI: 82–91); mean false-positive rate per scan was 3.8. Sensitivity across readers in detecting lesions ≥ 3 mm was statistically higher using CAD: 65% (95% CI: 61–69) versus 88% (95% CI: 86–91, $p<0.01$). Reading time increased by 11% using CAD (296 s vs. 329 s; $p<0.05$).

Conclusion In patients with extra-thoracic malignancies, CAD-assisted reading improves detection of ≥ 3-mm lung nodules on CT, slightly increasing reading time.

Key Points

- *CAD-assisted reading improves the detection of lung nodules compared with unassisted reading on CT scans of patients with primary extra-thoracic tumour, slightly increasing reading time.*
- *Cloud-based CAD systems may represent a cost-effective solution since CAD results can be reviewed while a separated cloud back-end is taking care of computations.*
- *Early identification of lung nodules by CAD-assisted interpretation of CT scans in patients with extra-thoracic primary tumours is of paramount importance as it could anticipate surgery and extend patient life expectancy.*

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-018-5528-6) contains supplementary material, which is available to authorized users.

✉ Lorenzo Vassallo
lorenzo.vassallo@ircc.it

[1] Department of Radiology at Candiolo Cancer Institute-FPO, IRCCS, Strada Provinciale 142 km 3.95, 10060 Candiolo, Turin, Italy

[2] Department of Surgical Sciences, University of Turin, A.O.U. Città della Salute e della Scienza, Via Genova 3, 10126 Turin, Italy

[3] Department of Applied Science and Technology, Polytechnic University of Turin, Corso Duca Degli Abruzzi 24, 10129 Turin, Italy

[4] Department of Radiation Oncology (MAASTRO), GROW School for Oncology and Developmental Biology, Maastricht University Medical Centre, Dr Tanslaan 12, 6229 ET Maastricht, The Netherlands

[5] Medical Physics Unit, Candiolo Cancer Institute-FPO, IRCCS, Strada Provinciale 142 km 3.95, 10060 Candiolo, Turin, Italy

[6] Pisa Section of INFN, Largo Bruno Pontecorvo 3, 56127 Pisa, Italy

[7] Turin Section of INFN, Via Pietro Giuria 1, 10125 Turin, Italy

Keywords Tomography · Radiologists · Lung · Metastases · Neoplasm

Abbreviations

CAD Computer-aided detection
CT Computed tomography
MIP Maximum intensity projection

Introduction

The lung is the second most common site of metastases from extra-thoracic malignancies after the liver [1]. Approximately 1/3 of patients with extra-thoracic cancer will develop lung metastases during the disease course [2]. Patients with lung metastases detected early after their onset have a higher chance of radical treatment [3]. Pulmonary metastasectomy is feasible in patients with five or fewer metastases, who retain a good pulmonary function if their primary tumour is controlled [4]. Five-year survival rates in patients undergoing surgery for lung metastases range from 30% to 62% [3, 5, 6]. Computed tomography (CT) is the state of the art modality for the detection of pulmonary metastases. Unfortunately, the CT detection rate of lung metastases for single and double reading is only 50% and 63%, respectively [7], and patients present with approximately one-third more lung nodules during thoracotomy [8]. The small dimension of most lung nodules, location proximal to large vessels and the large number of CT slices to be reviewed have been suggested as possible causes of radiologist misses [9].

Computer-aided detection (CAD) algorithms have been developed to support radiologists in the detection of pulmonary nodules [10–15]. CAD is usually adopted as a second reader, after radiologists have reviewed CT scans unassisted [13]. Retrospective studies have shown that CAD improves the detection of lung nodules in individuals undergoing low-dose CT for lung screening from 78.1% to 96.7% [14]. In the same context, CAD increased inter-observer agreement from 77% to 84% [15]. However, to our knowledge, the performance of CAD in the detection of lung metastasis is largely unknown since data are available from only one small retrospective study [16].

Despite the benefits in terms of overall detection performance, CAD systems have not been adopted extensively into clinical practice because of several drawbacks: the high fixed cost of software licenses, the need for a dedicated hardware and their rapid obsolescence [17]. Furthermore, the computational needs of CAD algorithms, depending on their complexity, often required powerful and expensive hardware [17].

For the above-mentioned reasons, CAD systems do not represent a cost-effective and time-efficient solution for most clinical facilities.

The 'Software-As-Service' approach, based on Cloud Computing resources accessible via secure Web protocols, makes CAD functionality available to users without requiring any dedicated hardware installation at the local facility [17, 18].

The main aim of this study was to compare unassisted and CAD-assisted detection and time efficiency of radiologists in reporting lung nodules on CT scans taken from patients with extra-thoracic malignancies using a Cloud-based system.

Materials and methods

Study design and patient population

Patients with an extra-thoracic primary tumour who consecutively underwent CT either for staging, restaging or follow-up of extra-thoracic malignancies at our Institute between September 2015 and March 2016 were enrolled in this retrospective observational single-centre study. Subjects with severe pulmonary fibrosis, diffuse bronchiectasis, extensive inflammatory consolidation and massive pleural effusion were excluded from the study as these conditions may mimic or hide lung nodules. Patients with more than ten nodules were also excluded from the study. While this is an arbitrary choice, in our opinion the inclusion of patients with a large number of nodules, occasionally in the hundreds, is unfeasible because of the long reporting times and the tediousness of the reporting process. Exclusion of patients with a high number of findings is a common procedure in CAD validation studies [19].

Approval by the local ethics committee was obtained prior to the beginning of the study and all patients signed written informed consent.

Computed tomography

CT scans were performed on a 128-detector row CT scanner (Somatom Definition Flash, Siemens Healthineers, Erlangen, Germany) according to the clinical practice of our institute. 209 chest and abdominal examinations were conducted using intravenous injection of iodinated contrast medium; the remaining 28 chest CT scans were performed without contrast administration. The following scanning protocol was adopted for contrast-enhanced CT: tube voltage, 140 kV; effective tube

current, 139/180 mAs with CARE dose; pitch, 0.9; detector configuration, 64.0×0.6 mm; iterative reconstruction algorithm SAFIRE strength 3. Conversely, unenhanced CT was performed as follows: tube voltage, 100 kV; effective tube current, 130 mAs with CARE dose; pitch, 1.2; detector configuration, 32.0×1.2 mm; iterative reconstruction algorithm SAFIRE strength 2. All images were reconstructed at 2 mm/1.7 mm using a kernel B70f filter, stored on the hospital PACS (Telemis), anonymised and then uploaded through the CAD WEB front-end (M5L lung CAD on-demand, INFN).

Study interpretation

The Cloud-based research prototype system used for this study (M5L lung CAD on-demand), the algorithms, their performances and the infrastructure have been described and validated in detail elsewhere [17, 18, 20]. Previously reported stand-alone sensitivity of the M5L lung CAD for lesions ≥ 3 mm was 80% with an average of four false-positive (FP) findings per scan [17]. Two experienced radiologists (reader 1= DR, reader 2= DC, 35 and 20 years of experience, respectively) and one resident radiologist (reader 3= LV, 3 years of experience) aware that patients had extra-thoracic malignancies, but blinded to disease prevalence, reported all CT examinations. Readers reported using CAD as the second reader as follows: First, the list of anonymised CT examinations was interpreted by each of the three radiologists unassisted using the hospital PACS. Exams were reviewed using maximum intensity projection (MIP) volume-rendering technique with a slice thickness of 5 mm [21]. Readers measured the longest diameter of each nodule using a lung window on the plane where the lesion was more conspicuous, provided nodule coordinates on the three planes (x, y, z) and described nodule structure (calcified, solid, part-solid and sub-solid) according to Fleischner Society's glossary of terms for thoracic imaging [22]. All information was stored on an on-line web-form and no additional change was allowed after validation.

Once unassisted reading was completed, CAD findings became accessible through a dedicated web-form (Fig. 1). CAD marks were automatically compared with those of unassisted reading. The software was programmed so that the unassisted and CAD-assisted findings were matched if the Euclidean distance between the 3D centres of the findings was smaller than or equal to the nodule diameter. Through the web-form, the radiologist could then classify unmatched findings as: FPs when prompts were not nodules, true-positives (TPs) if nodules were ≥ 3mm or irrelevant (IRR) if lesions were < 3 mm [19]. TPs were automatically added to the original unassisted report and the final findings were stored in a database while FPs and IRR findings were ignored. Average exam processing time was 19 min, and up to 14 cases could be processed in parallel [17, 18].

Reference standard

Two additional radiologists (GC, AM) not involved in the reading process and with more than 15 years of experience in reporting thoracic CT, used a plug-in specifically developed for the Osirix viewer [23] to review all the nodules ≥ 3 mm, annotated by at least one of the three aforementioned radiologists and/or by the CAD, stored in the database. The panel members in consensus reached a final decision and assessed whether each finding was a nodule (TP finding) or not (FP or IRR finding). In equivocal cases prior or follow-up CT examinations could be reviewed to support the reconciliation process. In order to classify the lesions according to their spatial location, the two additional radiologists classified lesions that abut a first-degree bronchus or blood vessel as central, lesions distant up to 2 cm from chest wall as peripheral and the remaining TP findings as intermediate [24]. Nodules were also classified according to their longest diameter in small (3–5 mm), intermediate (6–9 mm) and large (≥10 mm). Patients with no TP findings were classified as negative.

Reading time

Reading time was assessed as follows. First, an operator not involved in the reading process (A.T.) recorded the time of unassisted reporting for each reader. During unassisted reporting the reader was allowed to search freely for lesions on the institutional PACS system. Lesion coordinates were recorded as described earlier. Once unassisted reading was completed, CAD findings became accessible through a dedicated web-form and the additional time required for the review of CAD marks was also recorded.

Statistical analysis

The primary endpoint of the study was to compare the per-lesion detection rate of the CAD second reader paradigm with that of unassisted reading. Detection rate was defined as the ratio between the sum of TP findings of unassisted reading and of the CAD second reader paradigm, respectively, over the total number of TP lesions at the reference standard. Findings were stratified according to reader, size, location and nodule structure. Additional secondary endpoints were to compare per patient sensitivity, specificity and reading time of CAD second reader paradigm with that of unassisted reading. Per-patient sensitivity was defined as the ratio between the number of patients with at least one TP nodule at unassisted reading and at CAD second reader, respectively, over the total number of patients with at least one positive finding. Per-patient specificity was defined as the ratio between the number of TN patients at unassisted and CAD second reader, respectively, over the total number of negative patients.

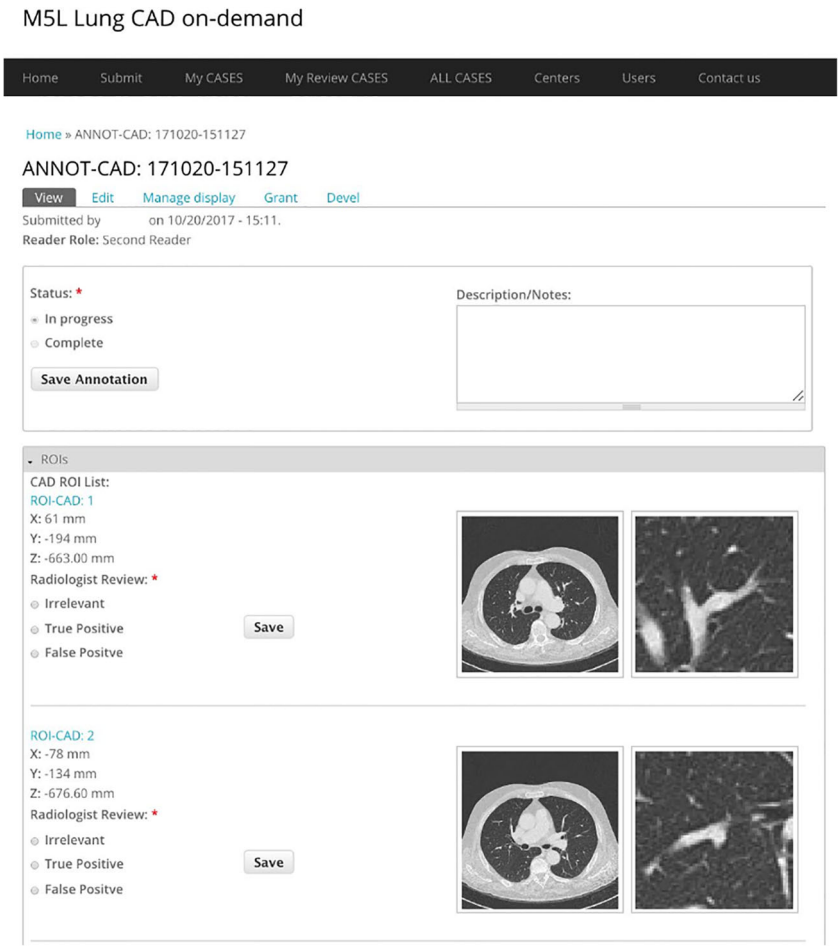

Fig. 1 Example of the Web and Cloud-based computer-aided detection interface

Per-patient and per-lesion detection rates of unassisted and CAD second reader paradigm were compared using McNemar's test. All statistical tests were 2-sided and statistical significance was set at $p \leq 0.05$. All confidence intervals (CIs) were reported at the 95% level using Microsoft Excel (Microsoft, Redmond, WA, USA, version 2016) and MedCalc (MedCalc Software, Ostend, Belgium, version 17.6). Bonferroni correction was applied as required for multi-observer studies. Confidence intervals and hypothesis testing was corrected as prescribed by Bonferroni, choosing significance at a threshold p-value < 0.05. The p-value was then divided by the number of observers (n=3) to obtain the final threshold for significance. Differences in sensitivity and specificity between readers with and without CAD were assessed using the 'N-1' chi-squared test [25, 26].

Results

The flowchart in Fig. 2 details the selection process for the study. The number of patients included into the study was 225, of whom 115 were men (mean age 60 years, range 21–90 years). At least one nodule was detected in 75 patients.

The most common primary tumours were: colorectal carcinoma (54/225, 24%), extra-thoracic sarcoma (25/225, 11%), melanoma (22/225, 10%) and breast cancer (21/225, 9%); other less common tumours combined accounted for the remaining 46% of cases (103/225). A total of 215 lung nodules with a diameter between 3 and 28 mm (median 8.5 mm) were identified at the reference standard. Seventy-seven percent (166/215) of nodules were small size, followed by intermediate (19%, 40/215) and large (4%, 9/215). Nodules were located at the periphery of the lungs in 62% of cases (133/215), while in 11% (23/215) they were positioned centrally. Most nodules were solid (67%, 143/215). Part-solid, sub-solid and calcified nodules accounted for 15% (33/215), 12% (25/215) and 6% (14/215) of lesions.

Stand-alone CAD performance

CAD stand-alone performances are summarized in Table 1. Overall per-lesion CAD sensitivity was 85% (183/215, 95% CI: 82–91). No significant sensitivity differences were observed comparing lesions of different sizes, location and structure. Of the 32 nodules missed by CAD, four (13%) were ground glass opacities and 11 (34%) were small (3–

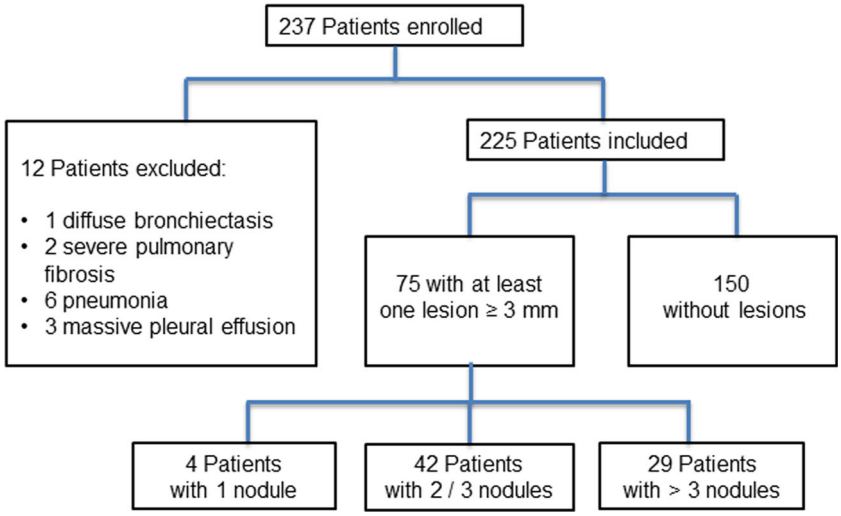

Fig. 2 Flowchart of the study population

5 mm) nodules in contact with lung vessels. The mean number of FP detections per scan was 3.8 (range 3.6–4.2). CAD FPs were mainly attributed to prominent ribs, vertebral osteophytes and superposition of vascular structures (Fig. 3a and 3b).

Per-lesion analysis

Results of per-lesion analysis are shown in Table 1. Across all readers the overall sensitivity of unassisted reading was 65% (421/645, 95% CI: 61–69) versus a sensitivity of 88% (570/645, 95% CI: 86–91) for CAD-assisted reading (p<0.01). Mean per-lesion sensitivity of CAD-assisted reading was significantly higher for all three readers compared to unassisted read. The number of false-negatives (FNs) was 74 for reader 1, 65 for reader 2 and 85 for reader 3. Results of sensitivity for each reader, according to lesion size, location and type, are presented in the Electronic Supplemental Material. The number of FP findings of unassisted reading was 19 for reader 1, 30 for reader 2 and 41 for reader 3. After reviewing CAD marks the number of FPs increased to 23 for reader 1, 36 for reader 2 and 49 for reader 3.

Mean sensitivity of unassisted and CAD-assisted reading for the detection of small nodules was, respectively, 61% (304/498, 95% CI: 57–65) and 91% (451/498, 95% CI: 88–93; $p < 0.01$) while the sensitivity for intermediate size nodules was, respectively, 77% (92/120, 95% CI: 68–84) and 85% (102/120, 95% CI: 77–91). Compared to unassisted reading, the CAD-assisted paradigm improved the detection of sub-solid nodules from 61% (46/75) to 91% (68/75;

Table 1 Per-lesion sensitivity of computer-aided detection (CAD) stand-alone, of unassisted and CAD-assisted interpretation across all readers for lesion size, location and structure

Nodule	CAD stand-alone	Unassisted reading		CAD-assisted reading		
Size	Sensitivity (%)	Sensitivity (%)	95% CI	Sensitivity (%)	95% CI	p-value
3–5 mm	84% (139/166)	61% (304/498)	57–65	91% (451/498)	88–93	< 0.01
6–9 mm	88% (35/40)	77% (92/120)	68–84	85% (102/120)	77–91	0.8
≥.10	100% (9/9)	93% (25/27)	76–99	100% (27/27)	87–100	0.16
Location						
Central	87% (20/23)	57% (36/69)	44–68	89% (61/69)	78–95	0.02
Intermediate	86% (50/59)	63% (112/177)	56–70	89% (158/177)	84–93	0.04
Sub-pleural	86% (114/133)	68% (270/399)	63–72	88% (351/399)	84–91	0.03
Structure						
Solid	92% (132/143)	65% (279/429)	60–70	88% (378/429)	85–91	< 0.01
Part solid	88% (29/33)	63% (62/99)	52–72	87% (86/99)	79–93	0.03
Sub solid	84% (21/25)	61% (46/75)	49–72	91% (68/75)	82–96	0.02
Calcified	93% (13/14)	81% (34/42)	66–91	81% (34/42)	66–91	0.9
Overall	85% (183/215)	65% (421/645)	61–69	88% (570/645)	86–91	< 0.01

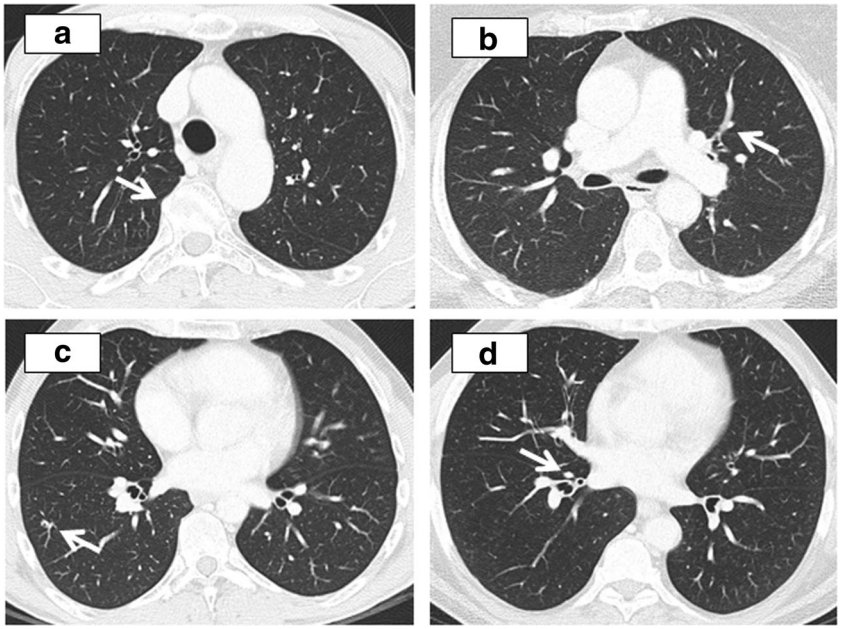

Fig. 3 Examples of the common computer-aided detection (CAD) findings. (**a**) CAD false-positive mark due to a prominent rib-vertebral joint. (**b**) CAD false-positive mark due to superposition of vascular structures. (**c**) Sub-solid nodule near to a small vessel at the periphery of the right lung detected only at CAD-assisted interpretation. (**d**) Small central nodule adjacent to a vessel detected only with the aid of CAD

p=0.02), of part-solid from 63% (62/99) to 87% (86/99; p<0.01) and of solid nodules from 65% (279/429) to 88% (378/429; p<0.01). Finally, mean sensitivity for central nodules was 57% (39/69, 95% CI: 44–68) and 89% (61/69, 95% CI: 78–95), respectively, for unassisted and CAD-assisted reading (p=0.02) (Fig. 3c and d).

Per-patient analysis

Results of the per-patient sensitivity analysis are shown in Table 2. Across all readers overall per-patient mean sensitivity of unassisted and CAD-assisted reading for the detection of nodules of all sizes was 75% (169/225 95% CI: 71–79) and 82% (186/225 95% CI: 79–86; p<0.01), respectively. Mean per-patient sensitivity of CAD-assisted reading was significantly higher for all three readers compared to unassisted reading. There were no differences in sensitivity between readers with and without CAD.

Results of the per-patient specificity analysis are shown in Table 3. Across all readers overall per-patient mean specificity of unassisted and CAD-assisted reading for the detection of nodules of all size was respectively 85% (382/450 95% CI: 81–88) and 82% (369/450 95% CI: 78–85; p=0.24). No differences in specificity were observed between unassisted and CAD-assisted reading for individual readers. Considering differences in specificity among readers, we found that reader 1 had a higher specificity than reader 3 (p=0.009) during the unassisted reading, while reader 1 had higher specificity than reader 2 (p=0.003) and reader 3 (p=0.002), when reporting with CAD.

Reading time and inter-observer variability

Table 4 shows reporting times for all the readers (and their mean) for both unassisted and assisted reading. With CAD reporting time increased on average by 11% (329 s CAD assisted vs. 296 s of unassisted read; p<0.05).

Table 2 Per-patient reader sensitivity of unassisted and computer-aided detection (CAD)-assisted interpretation

	Unassisted reading		CAD-assisted reading		
	Sensitivity	95% CI	Sensitivity	95% CI	p-value
Reader 1	75% (56/75)	68–81	83% (62/75)	76–88	0.04
Reader 2	80% (60/75)	73–85	85% (64/75)	79–90	0.045
Reader 3	70% (53/75)	63–77	80% (60/75)	73–85	0.02
Overall	75% (169/225)	71–79	82% (186/225)	79–86	0.01

Table 3 Per-patient reader specificity of unassisted and computer-aided detection (CAD)-assisted interpretation

	Unassisted reading		CAD-assisted reading		
	Specificity	95% CI	Specificity	95% CI	p-value
Reader 1	90% (136/150)	81–88	89% (133/150)	78–85	0.57
Reader 2	85% (127/150)	78–90	76% (114/150)	69–82	0.06
Reader 3	79% (119/150)	72–85	75% (113/150)	68–82	0.41
Overall	85% (382/450)	71–79	82% (369/450)	79–86	0.24

Table 4 Interpretation time of unassisted and computer-aided detection (CAD)-assisted reading

Reader	Unassisted reading Average time (time ± SD)	Review of CAD prompts Average time (time ± SD)	Total reading time Average time (time ± SD)	Percentage increase
Reader 1	(295 ± 85) s	(29 ± 10) s	(324 ± 97) s	+9.8%
Reader 2	(303 ± 70) s	(33 ± 12) s	(336 ± 90) s	+11%
Reader 3	(292 ± 72) s	(35 ± 13) s	(327 ± 75) s	+12%
Overall	(296 ± 80) s	(33 ± 12) s	(329 ± 83) s	+11%

Discussion

Early identification of metastatic lung progression is of paramount importance to identify patients with limited disease who potentially could benefit from a more aggressive and multidisciplinary approach [3–6]. In this study, we have shown that interpretation with CAD improves the overall sensitivity of radiologists by 23% in the detection of lung nodules with a size of 3 mm or more in patients with primary extra-thoracic tumours undergoing routine CT staging. To our knowledge, only one previous article tested a CAD paradigm in the detection of lung nodules in patients undergoing routine staging for extra-thoracic tumour, using a different reading paradigm where CAD was the first reader [16]. In the latter, CAD detected 64 of the 91 (70%) lung nodules in a total of 18 patients. All 27 lesions missed by CAD were in the sub-pleural layer. The specific algorithms implemented in our CAD system allowed the detection of 85% of sub-pleural nodules, 20% more than unassisted interpretation.

As already shown in screening series [19, 27–33], a higher sensitivity gain was observed with CAD for 3 to 5mm nodules (89% vs. 58% of unassisted readings), for sub-solid lesions (91% vs. 61% of unassisted readings) and for centrally located nodules (89% vs. 57% of unassisted readings). Central lesions are more likely to be overlooked by radiologists because of their adherence to vascular or bronchial structures, which can totally overlap the finding [19, 29].

It is worthwhile to explore differences between the performance of CAD in low-dose screening CT scans and in routine staging CT studies. Of note, the stand-alone performance of this research CAD system was higher (88% vs. 80%) in the present study than in a previous validation study on the LIDC-IDRI lung screening dataset, which also included low-dose scans [18]. Results of this study show the capability of this CAD system to achieve good performances also on external datasets without optimisation of the processing algorithms. The performance of CAD/reader could be affected by scanning protocols and, indeed, studies targeting early detection of bronchopulmonary carcinoma in low-dose lung CTs scans have shown improvement of the performance of CAD in the detection of pulmonary nodules, but not to the same extent as in our study [7, 14, 28–34]. According to a recent study on a screening population, CAD sensitivity decreased by 9% when cutting the dose to 50% of the original one, and the number of FPs increased significantly [35]. The performance of CAD could also be influenced by other factors such as CAD stand-alone sensitivity, the CAD algorithm itself, its operating point, reader experience, lesion conspicuity and inclusion criteria.

When assessing performance of a CAD system, the number of FP markings should also be taken into consideration. In our study, the mean number of FP detections per scan was 3.8, comparable to that of other studies [33, 34]. FP CAD prompts could affect reader specificity negatively by erroneously inducing the radiologist to point to lesions that are not nodules. However, in this study no significant difference was observed between the specificity of unassisted and CAD-assisted reading. The FP rate of this CAD system could be reduced by retraining the classifier in order to recognize the most common type of FPs, i.e. rib-vertebral joints and superposition of vascular structures (Fig. 3a and b).

One of the biggest concerns of radiologists when using CAD systems is the increased interpretation time, due both to the time required for post-processing of CT data and to the reporting scheme, since exams are read first unassisted and then reviewed with CAD. In this study processing of CT data was fully automated, requiring no additional doctor-time and therefore reading time increased on average only by 33 s. To make the CAD pipeline cost-efficient and available without software and hardware installation, a Software-as-Service approach was adopted where a web front-end handled image upload and annotation by radiologists and where all computing and storage resources were in a cloud infrastructure [17]. The time efficiency of this CAD system was improved compared to the performance of CAD as second reader in studies that did not adopt an automated matching system [27, 30, 36].

This study has some limitations. First, the reference standard for defining the presence of a pulmonary nodule was a consensus panel of only two experienced radiologists, who reviewed all nodules detected by the three readers and/or by CAD system. Surgical confirmation of the findings was not available. Potentially this process could have determined an underestimation of the number of TP nodules. Nevertheless, the hypothetical lower number of TP would have more likely affected the overall sensitivity of the two reading paradigms,

not the differences between CAD-assisted and non-assisted reading, which is the main aim of this study. Other studies, which used a consensus panel of radiologists as the reference standard, included a larger number of readers [28, 34–38]. However, reviewers in the consensus panel of this study could examine follow-up examinations of the same patients when deemed necessary. Second, this study was retrospective, and it may not necessarily reflect a real-life workflow. Third, the nature of the detected nodules was not assessed as in the specific clinical setting we considered (advanced extra-thoracic cancer) most patients did not undergo biopsy and/or surgery. However, the aim of this study was to assess the sensitivity in the detection of lung nodules in a well-defined, largely unexplored group of patients. Finally, the study had only three readers, including only one inexperienced reader, limiting the measure of inter-observer variability.

In conclusion, in this study we clinically validated a Cloud-based CAD system for the detection of lung nodules on CT scans of patients with extra-thoracic malignancies, showing sensitivity improvement in the CAD-assisted reading and only a limited increase of the reading time.

Funding The authors state that this work has not received any funding.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Alberto Traverso.

Conflict of interest The authors of this article declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry One of the authors has significant statistical expertise.

Informed consent Written informed consent was obtained from all subjects (patients) in this study.

Ethical approval Institutional Review Board approval was obtained.

Methodology
- retrospective
- observational
- performed at one institution

References

1. Mohammed TL, Chowdhry A, Reddy GP et al (2011) ACR Appropriateness Criteria® screening for pulmonary metastases. J Thorac Imaging. 26(1):W1–W3
2. Hornbeck K, Ravn J, Steinbruchel DA (2011) Outcome after pulmonary metastasectomy: analysis of 5 years consecutive surgical resections 2002–2006. J Thorac Oncol 6(10):1733–1740
3. Casiraghi M, De Pas T, Maisonneuve P et al (2011) A 10-year single-center experience on 708 lung metastasectomies: the evidence of the 'international registry of lung metastases'. J Thorac Oncol 6(8):1373–1378
4. Martini N, McCormack PM (1998) Evolution of the surgical management of pulmonary metastases. Chest Surg Clin N Am 8:13–27
5. Corona-Cruz JF, Domínguez-Parra LM, Saavedra-Pérez D et al (2012) Lung metastasectomy: long-term outcomes in an 18-year cohort from a single center. Surg Oncol 21(3):237–244
6. Okumura T, Boku N, Hishida T et al (2017) Surgical outcome and prognostic stratification for pulmonary metastasis from colorectal cancer. Ann Thorac Surg 104(3):979–987
7. Rubin GD, Lyo JK, Paik DS et al (2005) Pulmonary nodules on multi-detector row CT scans: performance comparison of radiologists and computer-aided detection. Radiology 234(1):274–283
8. Cerfolio RJ, McCarty TP, Bryant AS (2009) Non-imaged pulmonary nodules discovered during thoracotomy for metastasectomy by lung palpation. Eur J Cardiothorac Surg 35(5):786–791
9. Cerfolio RJ, Bryant AS, McCarty TP, Minnich DJ (2011) A prospective study to determine the incidence of non-imaged malignant pulmonary nodules in patients who undergo metastasectomy by thoracotomy with lung palpation. Ann Thorac Surg. 91(6):1696–1700
10. Ross 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(3): 549–557
11. Firmino M, Angelo G, Morais H, Dantas MR, Valentim R (2016) Computer-aided detection (CADe) and diagnosis (CADx) system for lung cancer with likelihood of malignancy. Biomed Eng Online 6 15:2
12. Parveen SS, Kavitha C (2012) A review on computer aided detection and diagnosis of lung cancer nodules. International Journal of Computers & Technology 3(3):393–400
13. White CS, Pugatch R, Koonce T, Rust SW, Dharaiya E (2008) Lung nodule CAD software as a second reader: a multicenter study. Acad Radiol 15(3):326–333
14. Zhao Y, de Bock GH, Vliegenthart R et al (2012) Performance of computer-aided detection of pulmonary nodules in low-dose CT: comparison with double reading by nodule volume. Eur Radiol. 22(10):2076–2084
15. Jeon KN, Goo JM, Lee CH et al (2012) Computer-aided nodule detection and volumetry to reduce variability between radiologists in the interpretation of lung nodules at low-dose screening computed tomography. Invest Radiol. 47(8):457–461
16. Schramm A, Wormanns D, Leschber G, Merk J (2011) Reliability of a computer-aided detection system in detecting lung metastases compared to manual palpation during surgery. Interact Cardiovasc Thorac Surg. 12(1):20–23
17. Traverso A, Cerello AM et al (2015) M5L: A web-based Computer Aided Detection system for automated search of lung nodules in thoracic Computed Tomography scans. Computational Vision and Medical Imaging Processing V:193–199
18. Torres EL, Fiorina E, Pennazio F et al (2015) Large scale validation of the M5L lung CAD on heterogeneous CT datasets. Med Phys. 42(4):1477–1489
19. Sahiner B, Heang-Ping C, Hadjiiski LM et al (2009) Effect of CAD on Radiologists' detection of Lung Nodules on Thoracic CT scans: Analysis of an observer performance study by nodule size. Acad. Radiol. 16(12):1518–1530
20. Retico A, Delogu P, Fantacci ME, Gori I, Preite Martinez A (2008) Lung nodule detection in low-dose and thin-slice computed tomography. Comput Biol Med. 38(4):525–534
21. James F, Gruden JF, Ouanounou S, Tigges S, Norris SD (2002) Incremental Benefit of Maximum-Intensity-Projection Images on Observer Detection of Small Pulmonary Nodules Revealed by Multidetector CT. AJR Am J Roentgenol 179:149–157

22. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Muller NL, Remy J (2008) Fleischner Society: Glossary of Terms for Thoracic Imaging1 Radiology: Volume 246: Number 3-march.
23. Rosset A, Spadola L, Ratib OJ (2004) OsiriX: an open-source software for navigating in multidimensional DICOM images. Digit Imaging 17(3):205–216
24. Letourneau PA, Xiao L, Harting MT et al (2011) Location of pulmonary metastasis in pediatric osteosarcoma is predictive of outcome. J Pediatr Surg. 46(7):1333–1337
25. Campbell I (2007) Chi-squared and Fisher-Irwin tests of two-by-two tables with small sample recommendations. Stat Med 26:3661–3675
26. Richardson JTE (2011) The analysis of 2 x 2 contingency tables - Yet again. Stat Med 30:890
27. Beyer F, Zierott L, Fallenberg EM et al (2007) 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 17(11):2941–2947
28. Marten K, Grillhösl A, Seyfarth T, Obenauer S, Rummeny EJ, Engelke C (2005) Computer-assisted detection of pulmonary nodules: evaluation of diagnostic performance using an expert knowledge-based detection system with variable reconstruction slice thickness settings. Eur Radiol. 15(2):203–212
29. Awai K, Murao K, Ozawa A et al (2004) Pulmonary nodules at chest CT: effect of computer-aided diagnosis on radiologists' detection performance. Radiology 230:347–352
30. Goo JM (2011) A Computer-aided diagnosis for evaluating lung nodules on chest CT: the current status and perspective. Korean J Radiol 12(2):145–155
31. Brown MS, Goldin JG, Rogers S et al (2005) Computer-aided lung nodule detection in CT: results of large-scale observer test. Acad Radiol. 12:681–686
32. Das M, Muhlenbruch G, Heinen S et al (2008) Performance evaluation of a computer-aided detection algorithm for solid pulmonary nodules in low-dose and standard-dose MDCT chest examinations and its influence on radiologists. Br J Radiol. 81:841–847
33. Beigelman-Aubry C, Raffy P, Yang W, Castellino RA, Grenier PA (2007) Computed-aided detection of solid lung nodules on follow-up MDCT screening: evaluation of detection, tracking, and reading time. AJR Am J Roentgenol. 189:948–955
34. Godoy MCB, Kim TJ, White CS et al (2013) Benefit of Computer-Aided Detection Analysis for the Detection of Subsolid and Solid Lung Nodules on Thin- and Thick-Section CT. AJR Am J Roentgenol. 200:74–83
35. Young S, Lo P, Kim G et al (2017) The effect of radiation dose reduction on computer-aided detection (CAD) performance in a low-dose lung cancer screening population. Med Phys 44(4): 1337–1346
36. Matsumoto S, Ohno Y, Aoki T et al (2013) Computer-aided detection of lung nodules on multidetector CT in concurrent-reader and second-reader modes: a comparative study. Eur J Radiol 82(8): 1332–1337
37. Armato SG 3rd, McLennan G, Bidaut L et al (2011) The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans. Med Phys. 38(2):915–931
38. Jacobs C, van Rikxoort EM, Murphy K, Prokop M, Schaefer-Prokop CM, van Ginneken B (2016) Computer-aided detection of pulmonary nodules: a comparative study using the public LIDC/IDRI database. Eur Radiol 26(7):2139–2147