

# UCLA

## UCLA Previously Published Works

### Title

Integration and evaluation of chest X-ray artificial intelligence in clinical practice.

### Permalink

<https://escholarship.org/uc/item/9ms0t53j>

### Journal

Journal of Medical Imaging, 10(5)

### ISSN

2329-4302

### Authors

Wong, Koon-Pong

Wei, Cindy

Yaghmai, Nazanin

et al.

### Publication Date

2023-09-01

### DOI

10.1117/1.JMI.10.5.051805

Peer reviewed

# Integration and evaluation of chest X-ray artificial intelligence in clinical practice

Koon-Pong Wong<sup>1b, a</sup>, Suzanne Y. Homer,<sup>b</sup> Sindy H. Wei<sup>1b, b</sup>,  
Nazanin Yaghmai,<sup>b</sup> Oscar A. Estrada Paz,<sup>c</sup> Timothy J. Young,<sup>c</sup>  
Russell G. Buhr<sup>1b, c</sup>, Igor Barjaktarevic,<sup>c</sup> Liza Shrestha,<sup>a</sup> Morgan Daly<sup>1b, a</sup>,  
Jonathan Goldin,<sup>a</sup> Dieter R. Enzmann,<sup>a</sup> and Matthew S. Brown<sup>a,\*</sup>

<sup>a</sup>David Geffen School of Medicine at UCLA, Center for Computer Vision and Imaging Biomarkers, Department of Radiological Sciences, Los Angeles, California, United States

<sup>b</sup>David Geffen School of Medicine at UCLA, Acute Care Imaging Section, Department of Radiological Sciences, Los Angeles, California, United States

<sup>c</sup>David Geffen School of Medicine at UCLA, Division of Pulmonary, Critical Care and Sleep Medicine, Los Angeles, California, United States

## Abstract

**Purpose:** To integrate and evaluate an artificial intelligence (AI) system that assists in checking endotracheal tube (ETT) placement on chest x-rays (CXRs) in clinical practice.

**Approach:** In clinical use over 17 months, 214 CXR images were ordered to check ETT placement with AI assistance by intensive care unit (ICU) physicians. The system was built on the SimpleMind Cognitive AI platform and integrated into a clinical workflow. It automatically identified the ETT and checked its placement relative to the trachea and carina. The ETT overlay and misplacement alert messages generated by the AI system were compared with radiology reports as the reference. A survey study was also conducted to evaluate usefulness of the AI system in clinical practice.

**Results:** The alert messages indicating that either the ETT was misplaced or not detected had a positive predictive value of 42% (21/50) and negative predictive value of 98% (161/164) based on the radiology reports. In the survey, radiologist and ICU physician users indicated that they agreed with the AI outputs and that they were useful.

**Conclusions:** The AI system performance in real-world clinical use was comparable to that seen in previous experiments. Based on this and physician survey results, the system can be deployed more widely at our institution, using insights gained from this evaluation to make further algorithm improvements and quality assurance of the AI system.

© 2023 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.10.5.051805](https://doi.org/10.1117/1.JMI.10.5.051805)]

**Keywords:** chest x-ray; endotracheal tube; artificial intelligence; clinical translation; user survey.

Paper 23019SSR received Jan. 15, 2023; accepted for publication Apr. 3, 2023; published online Apr. 25, 2023.

## 1 Introduction

Despite the many papers published on artificial intelligence (AI) for radiology, there are very few systems in routine clinical use. Systems have limited experimental testing and very few undergo evaluation in real-world use, although there are publications indicating that AI performance degrades significantly. Little attention is given to workflow integration and few studies assess system effectiveness and user satisfaction in clinical practice.

---

\*Address all correspondence to Matthew S. Brown, [mbrown@mednet.ucla.edu](mailto:mbrown@mednet.ucla.edu)

In a previous study, we developed and tested experimentally an AI system that can assist in checking endotracheal tube (ETT) placement and issue alerts to physicians if the ETT tip is not correctly positioned.<sup>1</sup> Chest x-rays (CXRs) are used in the intensive care unit (ICU) to examine and monitor critically ill patients on life-supporting devices. ETTs are used to maintain airway patency and lung ventilation. The desired ETT tip position is within the mid trachea, approximately  $5 \pm 2$  cm above the carina, with the head of the patient in a neutral position.<sup>2</sup> Other recommendations state that the ETT tip position should be at least 2 cm above the carina so that it does not enter one of the bronchi.<sup>3</sup> There is a risk of inefficient ventilation and vocal cord injury if the ETT is too high,<sup>4</sup> and a risk of lung collapse, pneumothorax, and even death if the ETT tip is too low.<sup>5</sup> The American College of Radiology recommends that a CXR is performed following intubation to check tube placement,<sup>6</sup> with repositioning being required in about 15% of patients.<sup>5,7-10</sup> CXR is the gold standard for post-intubation checking and is required in clinical practice. Given that tertiary ICUs can generate hundreds of bedside CXRs every day to check for tube placement and the urgent need for intervention if the ETT is misplaced, it is not always practical to wait for radiology reads. ICU physicians often take a preliminary look at the CXR at the bedside and immediately adjust a misplaced tube. However, due to the low conspicuity of tubes, superposition of anatomy and medical devices, and suboptimal image quality and patient positioning of bedside CXRs, assessment of tube placement can be challenging, especially for non-radiologists.

The purpose of this work was to deploy CXR AI to assist in checking ETT placement in clinical practice and evaluate its real-world performance with user feedback to determine if broader usage is appropriate. We describe the integration and evaluation methods and present the results and insights gained.

## 2 Materials and Methods

### 2.1 Chest X-Ray AI System

Details of the CXR AI system have been described previously.<sup>1</sup> In brief, the AI system is built on an open source Cognitive AI framework for computer vision known as SimpleMind.<sup>11</sup> SimpleMind allows for embedding deep convolutional networks within a knowledge base represented using a semantic network. The knowledge base describes each object to be identified in the image, in this case the ETT, trachea, and carina. Each object is represented by a node in the semantic network. The node lists the object's expected characteristics relating to location, size, and shape. In SimpleMind, multiple software agents work together to segment the image into candidate regions, then select the best candidate based on object attributes described in the knowledge base. Thus, the ETT, trachea, and carina are automatically segmented and labeled.

Based on the literature,<sup>2,3</sup> the desired region in which ETT tip position should be located is inside the trachea and within  $5 \pm 2$  cm above the carina, with the head of the patient in a neutral position and at least 2 cm above the carina. The SimpleMind knowledge base defines a "safe zone" for the ETT tip position inside the trachea and within  $5 \pm 2$  cm above the carina as per the medical literature.<sup>2,3</sup> The safe zone is an image region computed automatically by the system after it segments the trachea and carina, allowing the system to check whether the ETT tip is within this zone. This approach provides explanation in terms of how misplacement is determined by the AI, aiding in user confidence and adoption.

The AI automatically generates CXR overlays, showing the ETT path and distance, measured in centimeters, from the ETT tip to the carina. It also displays one of the three possible ETT messages: (1) "Found" (ETT tip was determined to be in the safe zone), (2) "Position Alert" (ETT tip was not in the safe zone or the AI could not determine the safe zone), (3) "Not Found" (no ETT was detected by the AI).

Similar to our previous work,<sup>1</sup> we evaluated the AI performance using definitions of true positive (TP), false positive (FP), true negative (TN), and false negative (FN) as shown in Table 1. The actual placement was established based on the radiology report, i.e., the radiologist interpretation in clinical practice rather than measurement thresholds as in our previous paper.<sup>1</sup> When alerts were issued, follow-up CXRs were reviewed and radiology reports checked for any

**Table 1** Confusion matrix definitions of AI alert versus ETT placement.

ETT actual placement	AI ETT output	
	“Position Alert”/ “Not Found”	“Found” (ETT correctly placed)
Incorrect	TP	FN
Correct	FP	TN

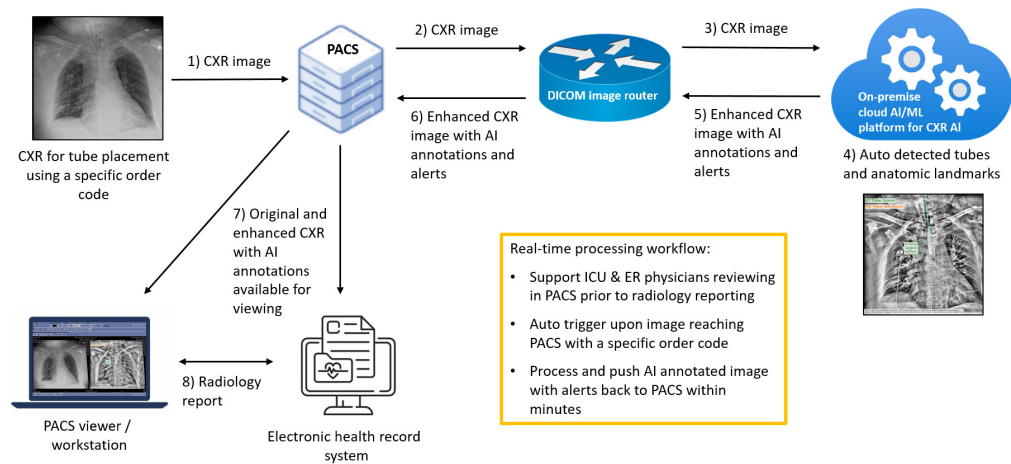
action taken (e.g., repositioning the tube). Thus, TP/TN references were determined by a combination of radiologist agreement provided in the radiology report and review of the follow-up CXR to determine whether the ETT was repositioned (for TP findings). Sensitivity to misplaced ETTs [sensitivity =  $TP/(TP+FN)$ ], positive predictive value [PPV =  $TP/(TP+FP)$ ] and negative predictive value [NPV =  $TN/(TN+FN)$ ] metrics were computed to give a sense of trustworthiness of the AI from a physician perspective.

In previous experimental testing on 285 CXRs with ETT, the PPV, NPV, and sensitivity to misplaced tubes were 42%, 99%, and 95%, respectively.<sup>1</sup> The AI system was designed to be highly sensitive to avoid missed alerts when ETTs were misplaced, thus higher NPV and lower PPV were considered sufficient for the system to be deployed in clinical practice and further evaluated as described in this paper.

## 2.2 Portable CXR Clinical Workflow

To plan the AI integration, we studied the clinical workflow for portable CXRs to check ETT placement in the ICU at our institution. After the CXR is acquired and pushed to the picture archiving and communication system (PACS), the workflow bifurcates into two streams, one for the ICU physician and another for the radiologist. ICU physicians usually have access to a PACS workstation or a zero-footprint DICOM viewer at the bedside and perform an initial review of the CXR within minutes to enable quick action if the ETT is misplaced. A radiologist will also review the CXR and issue a formal report, often calling the physician if they detect a misplaced tube. It is when the report is delayed or a miscommunication occurs that there is the potential for harm to the patient. Since both the ICU physician and the radiologist have access to PACS, it was determined that the AI outputs should be sent to the PACS. Another important clinical integration requirement was rapid and fully automated processing to ensure that AI outputs would be available to ICU physicians at the point of care during their initial CXR review.

A new cloud-based computing infrastructure was designed to integrate the CXR AI system, installed on a cloud-based AI/machine learning (ML) platform, with the relevant clinical systems, including the electronic health record (EHR), PACS, and a digital imaging and communications in medicine (DICOM) image router (see Fig. 1). The workflow requires that an ICU physician orders the CXR with AI processing using a specific order code in the EHR system. An image order shows on the worklist created and managed by the EHR system and remains until the order is marked as complete by the x-ray technologist who performs the CXR examination and pushes the DICOM image to PACS. The EHR system also automatically sends the worklist to a DICOM image router that is configured to pick up from PACS only the CXR scans ordered with the specific order code and push them to an on-premise Microsoft Azure AI/ML platform where the CXR AI system is deployed. The CXR AI processes the DICOM image, detects tubes and anatomic landmarks on the image, and generates an enhanced CXR image with overlays and an alert/informational message that will be picked up by the DICOM image router and subsequently pushed to PACS. The AI-generated image, created as a DICOM Secondary Capture image, has the same accession number and DICOM study instance unique identifier (UID) as the original CXR image, but a distinct series instance UID and service-object pair instance UID so that the images are appropriately associated in PACS. The original CXR and the AI-generated image showing the overlays of ETT path and distance measurement as well as the ETT message, are available in PACS and EHR systems for the radiologist and ICU physician to view on a PACS workstation or a zero-footprint DICOM viewer. The total turnaround time



**Fig. 1** Simplified diagram of integration of clinical information systems with an on-premises cloud AI/ML platform for running the CXR AI.

from CXR reaching PACS to the AI output being available in PACS is within 3 to 4 min. This turnaround time was considered sufficient by the ICU physicians who participated in the clinical AI workflow development.

Risk assessment was conducted by our institutional compliance office prior to deployment of the CXR AI system in the clinical environment. Standard software tests, including unit testing, functional testing, integration testing, acceptance testing, performance testing, and stress testing, were carried out to test, verify, and validate individual and integrated units/components to ensure they were operating correctly in a testing environment before installation in the production environment.

### 2.3 Image Routing Criteria

The CXR AI is currently deployed within our institution's clinical workflow for investigational use only as a quality improvement tool. A specific order code was set up for CXR with AI processing, and access to this order code was granted only to the ICU physicians involved in the AI evaluation. This provided a limited deployment on identifiable cases to be reviewed by them along with a selected pool of radiologists specializing in acute care imaging. From June 11, 2021 to November 3, 2022, 214 CXRs were ordered by ICU physicians through this specific order code for checking ETT placement with AI assistance.

### 2.4 Evaluation of AI Overlay and Alerts

The AI alerts and ETT overlay were evaluated against the findings in the radiology report in which the radiologists were asked to include the following statement using a smart text feature: "an investigational ETT AI overlay was available and was/was not consistent with my interpretation." If they indicated in the smart text that the AI overlay was not consistent with their interpretation, they were asked to include their observation as a probable cause so we could mine the data for future improvement of the AI model and for implementation of quality assurance. In cases where the AI overlay was not yet available when they dictated their initial radiology report, they were asked to add the smart text as an addendum to their report upon subsequent review.

### 2.5 User Survey

A simple survey was conducted to qualitatively evaluate the clinicians' experience in using the CXR AI in their clinical workflow. Several clinicians, including ICU physicians and radiologists from the acute care imaging section, of varying levels of experience in their specialties were invited to participate in this user evaluation. They were asked to provide rating and feedback on a questionnaire regarding the system outputs (SO) as well as usefulness and satisfaction (US)

System Outputs		1	2	3	4	5	
The system output agrees with my assessment	Strongly disagree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly agree
Tube annotations are helpful/appropriate	Strongly disagree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly agree
Informational/alert messages are helpful/appropriate	Strongly disagree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly agree
The system output increases my confidence	Strongly disagree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly agree

Usefulness and satisfaction		1	2	3	4	5	
It helps me be more effective	Strongly disagree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly agree
It helps me be more productive	Strongly disagree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly agree
It works the way I want it to work	Strongly disagree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly agree
I am satisfied with it	Strongly disagree	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Strongly agree

---

How many cases have you been reviewed using the CXR AI system?

> 5                       > 20                       > 50

> 10                       > 30

---

Please provide any feedback or comment about the system:

**Fig. 2** Questionnaire to collect user evaluations and experience in using the AI clinical workflow.

with the AI clinical application. Figure 2 shows the questionnaire used in this study. The questionnaire was developed based on a published usability assessment tool.<sup>12</sup> It contains eight items on a 5-point rating scale (1 = “strongly disagree;” 5 = “strongly agree”). Four items related to SO and four to US. Survey results included a summary of the number of respondents, their roles, and experience in their specialty, and the number of times they had used the AI in their workflow. Descriptive statistics on the survey ratings included mean, standard deviation (SD), median, minimum, maximum, skewness, and kurtosis. Spearman correlation coefficient was computed between the rating and the respondent’s years of experience in their specialty.

### 3 Results

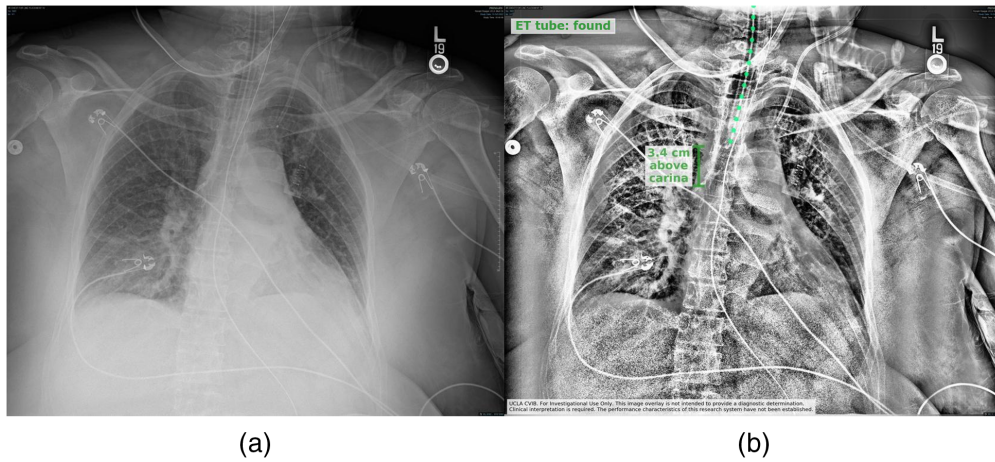
Table 2 shows a confusion matrix of AI alert versus actual ETT placement per radiology report. For the 214 CXR images processed, the CXR AI had a PPV of 42% [21/(21 + 29)] and an NPV of 98% [161/(161 + 3)] in clinical use. These clinical evaluation metrics are listed in Table 3 along with those from previous experimental testing for ease of comparison.

**Table 2** Confusion matrix of AI alert versus ETT placement.

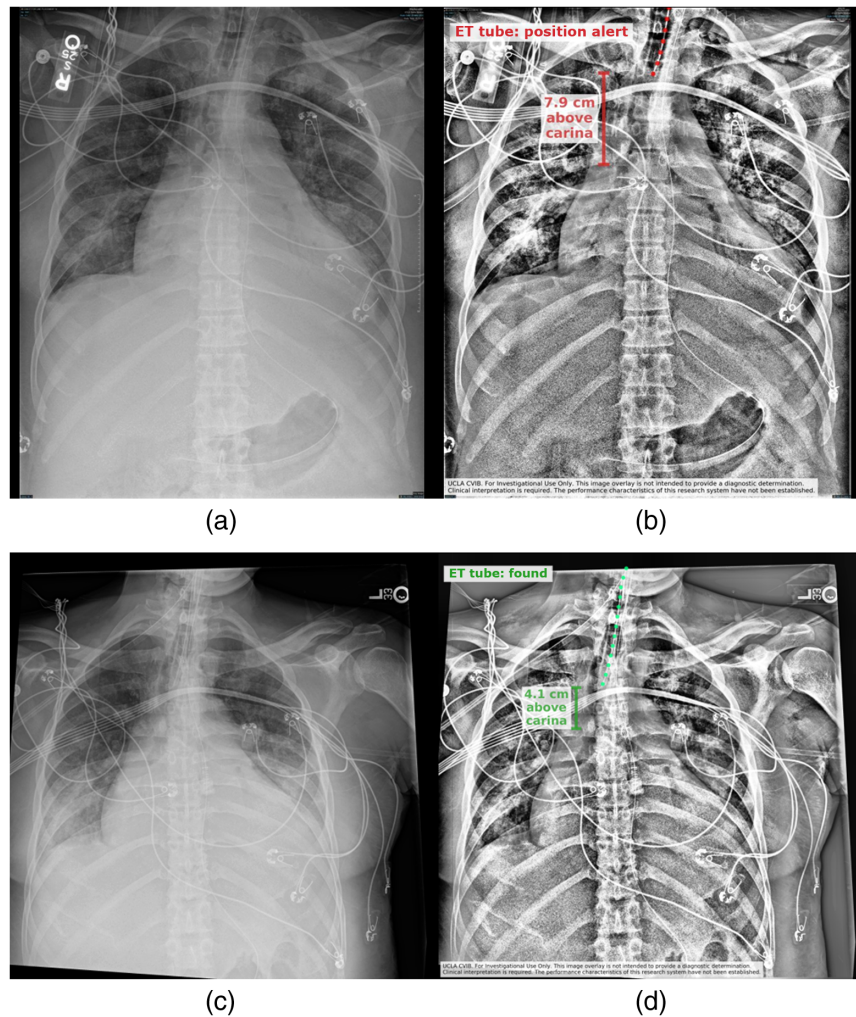
ETT actual placement	AI ETT output	
	“Position Alert”/ “Not Found”	“Found” (ETT correctly placed)
Incorrect	21	3
Correct	29	161

**Table 3** CXR AI performance metrics for previous experimental testing and current clinical evaluation.

	Number of CXRs	PPV	NPV	Sensitivity to misplaced ETTs
Previous experimental testing	285	42%	99%	95%
Clinical evaluation	214	42%	98%	88%



**Fig. 3** (a) Original CXR and (b) the AI output image for a TN case with a correctly placed ETT.



**Fig. 4** (a) Original CXR and (b) the AI output image for a TP case in which the ETT tip was placed too high. (c) The follow-up CXR and (d) the AI output. The ETT tip was repositioned to a proper location in the follow-up scan.

The ETT was determined clinically (based on radiology reports) to be correctly placed in 190 cases and misplaced in 24 cases (11%). Among the 190 correctly placed ETT cases, the system correctly called 161 TNs with 29 FPs. Figure 3 shows an example of a TN for a correctly placed ETT. Of the 29 FPs, 22 were unnecessary alerts for ETT misplacement caused mainly by the carina or the ETT tip being obscured by other devices, and 7 FPs resulted from the ETT not being detected due to partial/complete overlapping of the ETT with other wires.

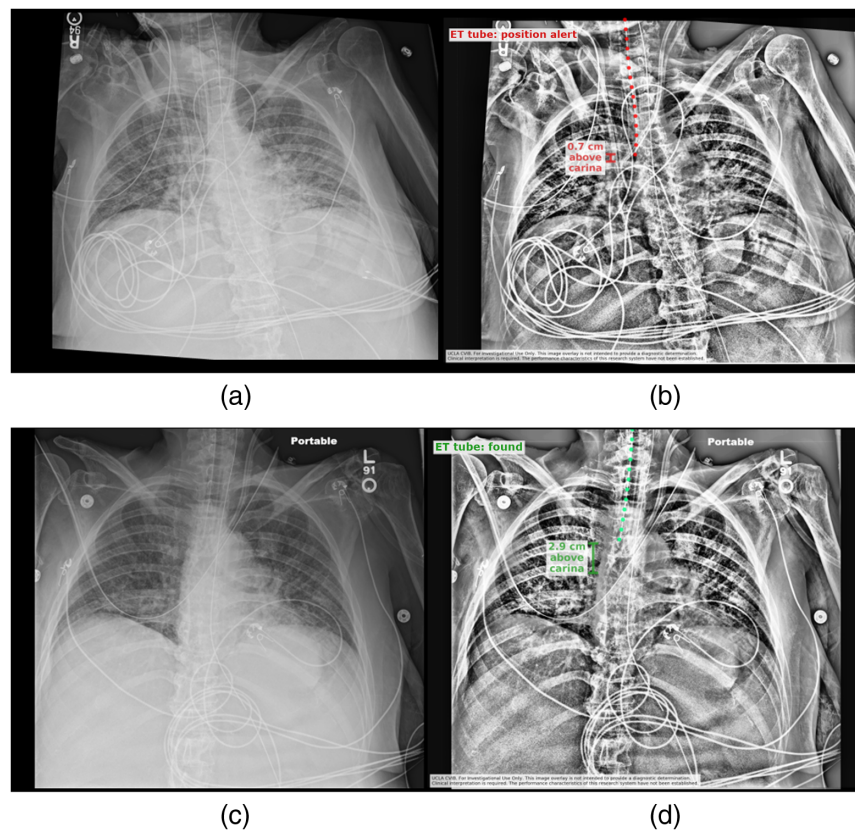
For the 24 cases with misplaced ETT, positive alerts (TPs) were generated by the AI in 21 cases (see examples in Figs. 4 and 5, also showing follow-up CXRs with corrected ETT positions). There were three FNs due to inaccurate carina and/or ETT tip localization. Figure 6 shows one such case in which the AI correctly predicted the ETT tip location, which was close to the carina, but the AI-predicted carina location was about a centimeter lower than the actual location, resulting in a FN output.

ETT tip-to-carina measurements were available for both AI and radiology report in 174 cases. Median and mean ( $\pm$ SD) of the absolute difference between radiologist's and AI's ETT tip-to-carina measurements were 0.6 cm and  $0.78 \pm 0.83$  cm, respectively.

For cases in which the radiologists indicated in the smart text that AI overlay was not consistent with their interpretation, their disagreement was mainly with the carina location detected by the AI being lower than the actual location. Visual inspection of the enhanced CXR images revealed that most of the time when this occurred the carina was obscured by overlying devices/wires in close proximity.

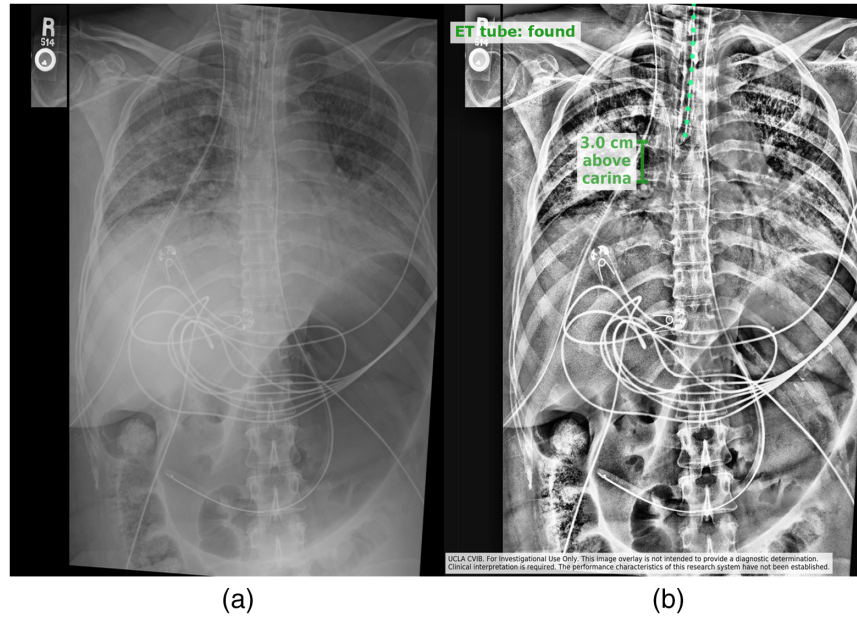
All of the 17 clinicians who used the CXR AI workflow were invited to participate in the user survey. Seven returned the questionnaire with ratings: three were board certified radiologists specializing in diagnostic radiology with 9 to 26 years of experience, and four were physicians board certified in internal medicine, having 1 to 10 years of experience in critical care medicine.

A label was assigned to each item for the ease of data analysis and was not shown in the questionnaire form. Table 4 summarizes the frequency, mean and median of the rating for all



**Fig. 5** (a) Original CXR and (b) the AI output image for a TP case in which the ETT tip was placed too low. (c) The follow-up CXR and (d) the AI output image. The ETT tip was repositioned to a proper location in the follow-up scan.





**Fig. 6** A false-negative case in which the AI-predicted carina location was incorrect. (a) Original CXR and (b) AI output image.

**Table 4** Frequency, mean, and median of rating for each survey question.

Label	System Outputs	Rating <sup>a</sup>					Mean <sup>b</sup>	Median
		1	2	3	4	5		
SO1	The SO agrees with my assessment	0	0	1	4	2	4.1 ± 0.7	4
SO2	Tube annotations are helpful/appropriate	0	0	2	3	2	4.0 ± 0.8	4
SO3	Informational/alert messages are helpful/appropriate	0	0	3	2	2	3.9 ± 0.9	4
SO4	The SO increases my confidence	0	0	2	3	2	4.0 ± 0.8	4

Label	Usefulness and Satisfaction	Rating <sup>a</sup>					Mean <sup>b</sup>	Median
		1	2	3	4	5		
US1	It helps me be more effective	0	0	3	4	0	3.6 ± 0.5	4
US2	It helps me be more productive	0	1	4	2	0	3.1 ± 0.7	3
US3	It works the way I want it to work	0	0	3	4	0	3.6 ± 0.5	4
US4 <sup>c</sup>	I am satisfied with it	0	0	2	4	0	3.7 ± 0.5	4

	Frequency				
	> 5	> 10	> 20	> 30	> 50
How many cases have you been reviewed using the CXR AI system	2	0	1	2	2

<sup>a</sup>1 = “Strongly disagree;” 5 = “Strongly agree.”

<sup>b</sup>Mean ± SD.

<sup>c</sup>N = 6.

**Table 5** Descriptive statistics of SO and US categories.

Category	Mean	SD	Minimum	Median	Maximum	Skewness	Kurtosis
<b>SO</b>	4	0.8	3	4	5	0	-1.3
<b>US</b>	3.5	0.6	2	4	4	-0.6	-0.6

items. One of the respondents did not provide a rating for US4. Most of the items had a median rating of 4 except US2 that had a median rating of 3. Five of the seven clinicians had reviewed over 20 CXR examinations with AI, and two of them had reviewed over 50. The overall rating of the SO and US was computed by averaging the item ratings in the corresponding category and the results are summarized in Table 5. The survey results were not associated with the clinicians' years of experience (Spearman correlation,  $P > 0.05$ ).

#### 4 Discussion

We previously did a classic ML train and test evaluation in a controlled experiment. This paper contributes to clinical translation of AI to radiology practice in describing: (1) integration of AI into the clinical workflow, (2) ongoing quality control and monitoring of AI performance metrics, and (3) evaluation of usefulness and user satisfaction.

The clinical evaluation showed good performance of the CXR AI system and consistency in PPV and NPV with previous experimental testing<sup>1</sup> (see Table 3). As described previously, the system was designed for higher NPV than PPV since the risk of not alerting on a misplaced ETT outweighs the user fatigue arising from false positive alerts. In the previous experimental testing, the ground truth of a misplaced ETT was derived by thresholding measurements of the distance from the ETT tip to the carina from annotations by a trained image analyst. In the clinical evaluation, cases of misplaced ETT were based on the radiologist interpretation as documented in their report. This may have contributed to a lower proportion of cases (24 of 214 cases) being labeled as misplaced in the clinical evaluation compared to the previous experimental testing (42 of 285 cases). The low number of cases with misplaced ETTs impacts our ability to reliably measure AI sensitivity. However, there was a decrease in the measured AI sensitivity to misplaced ETTs (88%) compared to in the previous experimental testing (95%). This may be due to differences in the annotated position of the carina in the training set, compared to the clinical assessment of carina position by the radiologists evaluating the system (see Fig. 6). During the clinical evaluation, the radiologists reported that the position of the carina indicated by the AI tended to be lower than their interpretation. These insights suggest that we need to focus on monitoring and potentially further model refinement on improving the AI detection of the carina location. This experience shows the general importance of monitoring AI performance continuously in clinical practice. In observing the consistency between AI performance in clinical practice and experimental testing in this study, it should be noted that both involved data from our institution. Ongoing performance monitoring would be necessary if the system were deployed at another institution as there is the potential for greater variation.

The user survey results (see Table 4) indicated overall agreement with the AI outputs and appropriateness of the alerts by both radiologists and ICU physicians. In terms of the usefulness of the system, user ratings suggest that while the AI does not save them time (increase productivity), it does increase their confidence and works the way they would expect AI to in their workflow. The variability among user responses suggests that larger surveys of more users are warranted. In clinical practice, ICU physicians do not record their interpretation of the CXR, so we could not evaluate any differences in diagnosis between the radiologists and ICU physicians.

Barriers to adoption of clinical decision support systems (CDSS) and best practices have been described by Scheepers-Hoeks et al.<sup>13</sup> It has been noted that few studies have examined the technical correctness and clinical appropriateness of alerts and proposed that a multidisciplinary expert panel is necessary to conduct this assessment on an ongoing basis in clinical

practice. The expert team should check for clinical relevance, usefulness of all the alerts and ensure that they can be acted on. In our study, the expert team consisted of ICU physicians, radiologists from the acute care imaging section, and engineers involved in the AI deployment. The team met weekly post deployment to monitor the effectiveness of the system and provide input on the correctness, usefulness, presentation, and frequency of alerts. The authors further advocate predictive value (PPV and NPV) as user-focused metrics in this assessment to maintain balance between clinical relevance and over-alerting, which are key drivers of adoption.<sup>13</sup> These were the primary metrics used in our evaluation and we have observed the importance of continuous maintenance post release of the AI decision support. Finally, it has been observed that there is a need to share the most effective practices in CDSS development and implementation and this work is such a contribution.

The clinical integration made use of next-generation technologies to achieve lower costs and higher scalability by deploying our AI software to an on-premises Microsoft Azure AI/ML platform. An on-premises cloud solution was used because it provides greater control and security over a public cloud environment. This is crucial in a healthcare setting involving personal health data. The type of integration that is appropriate depends very much on the clinical workflow. In our workflow, it was deemed best to output results to PACS. The degree of integration into the PACS is vendor dependent and it is possible that as PACS solutions evolve they may support more direct hosting of AI algorithms.

## 5 Conclusion

A system for checking ETT placement on CXR was built on the SimpleMind Cognitive AI platform and integrated into a clinical workflow and evaluated in routine practice. Performance metrics remained consistent with previous experimental testing and user survey results on the correctness and usefulness of the AI were positive. The system can be deployed more widely at our institution, using insights gained from this evaluation to make further algorithm improvements. The study demonstrated an approach to clinical deployment of AI and the importance of ongoing quality assurance and user feedback.

## Disclosures

The authors declare that there is no conflict of interest.

## Acknowledgments

This research received no particular grant from any funding agency in the public, private, or not-for-profit sectors.

## References

1. M. S. Brown et al., "Automated endotracheal tube placement check using semantically embedded deep neural networks," *Acad. Radiol.* **30**(3), 412–420 (2023).
2. L. R. Goodman et al., "Radiographic evaluation of endotracheal tube position," *Am. J. Roentgenol.* **127**(3), 433–434 (1976).
3. J. R. Roberts, M. Spadafora, and D. C. Cone, "Proper depth of placement of oral endotracheal tubes in adults prior to radiographic confirmation," *Acad. Emerg. Med.* **2**(1), 20–24 (1995).
4. M. D. Wiener et al., "Imaging of the intensive care unit patient," *Clin. Chest Med.* **12**(1), 169–198 (1991).
5. W. Brunel et al., "Assessment of routine chest roentgenograms and the physical examination to confirm endotracheal tube position," *Chest* **96**(5), 1043–1045 (1989).
6. A. T. Laroia et al., "ACR appropriateness Criteria® intensive care unit patients," *J. Am. Coll. Radiol.* **18**(5S), S62–S72 (2021).

7. P. Gray et al., "Value of postprocedural chest radiographs in the adult intensive care unit," *Crit. Care Med.* **20**(11), 1513 (1992).
8. P. Marik and M. Janower, "The impact of routine chest radiography on ICU management decisions: an observational study," *Am. J. Crit. Care* **6**(2), 95–98 (1997).
9. B. Trotman-Dickenson, "Radiology in the intensive care unit (Part I)," *J. Intensive Care Med.* **18**(4), 198–210 (2003).
10. M. C. B. Godoy et al., "Chest radiography in the ICU: Part 1, evaluation of airway, enteric, and pleural tubes," *Am. J. Roentgenol.* **198**(3), 563–571 (2012).
11. Y. Choi, M. W. Wahi-Anwar, and M. S. Brown, "SimpleMind: an open-source software environment that adds thinking to deep neural networks," *PLoS One* **18**(4), e0283587 (2023).
12. A. M. Lund, "Measuring usability with the USE questionnaire," *Usability Interface* **8**(2), 3–6 (2001).
13. A. M. J. Scheepers-Hoeks et al., "Strategy for development and pre-implementation validation of effective clinical decision support," *Eur. J. Hosp. Pharm.: Sci. Pract.* **20**(3), 155–160 (2013).

Biographies of the authors are not available.