NICE National Institute for Health and Care Excellence

Artificial Intelligence software for analysing chest X-ray images to identify suspected lung cancer : early value assessment

Section A: External Assessment Report - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
Annalise.ai	1	74	5	Per our correspondence with Frances, if other companies do not consistently lift their confidentiality for non-discounted prices, we would strongly prefer to have our pricing marked as commercial in confidence for the final report.	We are happy to be guided by NICE on this.
Annalise.ai	2	31 111	4.2 Appendix 2: table of excluded studies	Annalise's relevant studies have not been included in the report (neither in appendix 2 table of excluded studies nor section 4.2 'selected excluded' studies), suggesting that they were not found during the literature search. We have included summaries and links to the publications in the tables below as we feel we have relevant evidence for the scope of this EVA. Happy to discuss further.	The studies listed below were not provided by the company at the appropriate time in the EVA process and were not identified by the ERG searches. Nevertheless, none of the studies meet either the eligibility criteria listed in the protocol or the post hoc criteria developed in discussion with NICE. See below for individual studies.
Annalise.ai	3		**EVA Research Qs and Outcome measures taken from Table 1: eligibility criteria**	**Key question 1. What are the test accuracy and test failure rates of adjunct AI software to detect lung cancer on CXR?** **Sub-questions:** **1a. What is the test accuracy of adjunct AI software to detect lung nodules?** **1b. What is the concordance in lung nodule detection between radiology specialist with and without adjunct AI software** Test accuracy for the detection of lung cancer (sensitivity, specificity, positive predictive value, numbers of true positive, false positive, true negative, false negative results, number of lung cancers diagnosed) Test failures (rates, and data on inconclusive, indeterminate, and excluded samples, failure due to any other reason) Characteristics of discordant cancers cases Test accuracy for the detection of lung nodules Concordance in lung nodule detection between radiology specialist with and without adjunct AI software	See below

Artificial Intelligence software for analysing chest X-ray images to identify suspected lung cancer : early value assessment

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
Annalise.ai	4		Annalise evidence: Key, relevant outcomes	Retrospective multi-reader multi-case study. Published 2021, Lancet: https://www.thelancet.com/journals/landig/article/PIIS2589-7500(21)00106-0/fulltext#seccestitle10 Dataset: Australia, USA 486 cases contained one or more of the following findings (lung-cancer relevant): 'single_pulmonary_nodule', 'single_pulmonary_mass', 'cavitating_mass', 'cavitating_mass_internal_content', 'calcified_pulmonary_mass', 'multiple_pulmonary_masses', 'miliary' Statistically significant improvement in detection of lung-cancer relevant findings [**For each finding the Unassisted AUC and Assisted AUC** can be found in the Supplementary material, pages 7&8, table titled Primary endpoint performance metrics]	This study was not identified by the ERG searches and was not provided in the company submission. If identified, the study would have been excluded at the title and abstract stage as it has an ineligible population (inpatient, outpatient, and emergency settings). Therefore, it would not be listed in the list of excluded studies.
Annalise.ai	5		Ongoing studies (preliminary results, confidential)	[redacted]	This study was not identified by the ERG searches and was not provided in the company submission. If identified, it would have been excluded on the basis that it has an ineligible intervention (standalone AI)
Annalise.ai	6		**EVA Research Qs and Outcome measures taken from Table 1: eligibility criteria**	**Key question 2. What are the practical implications of adjunct AI software to detect lung cancer on CXR?[a]** Practical implications[a] (time to x-ray report, CT scan, diagnosis, turnaround time (image review to radiology report), acceptability of software to clinicians, impact on clinical decision-making, impact of false positives on workflow)	See below
Annalise.ai	7		Annalise evidence:	Prospective, multi-centre, real-world observational study, and post-study survey	This study was not identified by the ERG searches and was not provided in the company submission.

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
			Key, relevant outcomes	Published 2021, BMJ Open: https://bmjopen.bmj.com/content/11/12/e052902 (includes access to survey questions) Dataset: Australia. 2972 cases (consecutive CXRs from outpatient, inpatient and emergency settings) were reported by 11 radiologists over a period of 6 weeks. AI findings were displayed to the radiologist after their initial read of the case. For each CXR the reporting radiologists was asked to indicate whether the AI results led to changes. Key results: Viewing AI findings resulted in significant report changes for 3.1% of cases, a change in patient management for 1.4% of cases, and further imaging investigation for 1.0% of cases. Radiologists with fewer than 5 years consultant experience contributed 1347 cases, and indicated a rate of 5.0% for significant report change, 2.4% patient management change, and 1.5% recommendations for further imaging. Whether a patient was imaged as an inpatient or outpatient was not significantly associated with any change in report, patient management or imaging recommendation (p=0.358, p=0.572, p=0.326, respectively) Post-study survey results: (10 of 11 radiologists completed the survey) Seven (70%) participants felt that their reporting time was slightly worse, however, when asked how satisfied they were with their reporting time, seven (70%) indicated that they were satisfied. Nine out of 10 radiologists responded that their reporting accuracy was improved while using the CXR viewer, with 9 out of 10 (90%)	However, it would have been excluded as it includes an ineligible population (mix of outpatient, inpatient and emergency settings)

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				participants being satisfied with accuracy of the CXR model's findings. Nine radiologists (90%) demonstrated an improved attitude towards the use of the AI diagnostic viewer by the end of the study and 9 (90%) demonstrated an improved attitude towards AI in general. No radiologists reported a more negative attitude towards the CXR viewer or towards AI in general. Exploratory (feasibility) study of the impact of triage on reporting time and report turnaround times. Poster presentation, BIR. Non-peer reviewed: https://annalise.ai/resource/radiologist-reporting-productivity-benefits-from-ai-assisted-triage-of-cxr-studies-in-clinical-practice/ (Poster attached separately as a pdf) Dataset: 400 CXRs, 4 readers (100 unique CXRs per reader) Results: Consistent trend towards reduced reporting time for the AI-triaged worklists compared to the unsorted worklists. Average reporting time across all CXRs decreased from 55.8 to 48.5 seconds, while the reporting time for remarkable CXR studies decreased from 75.5 seconds to 66.8 seconds and reporting time for unremarkable CXR studies decreased from 35.5 seconds to 30.3 seconds.	This study was not identified by the ERG searches and was not provided in the company submission. However, it would have been excluded as it does not provide sufficient information for assessment of methodological quality, and reading time is not specific to lung cancer/nodules.
Annalise.ai	8		Ongoing studies (preliminary results, confidential)	[redacted]	Thank you for bringing these to our attention. We look forward to reading the results of these studies.

Artificial Intelligence software for analysing chest X-ray images to identify suspected lung cancer : early value assessment

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
Annalise.ai	9		**EVA Research Qs and Outcome measures taken from Table 1: eligibility criteria**	**Key question 3. What is the clinical effectiveness of adjunct AI software applied to CXR?** Mortality, morbidity, health-related quality of life	No studies are noted for this outcome. No action required.
Behold.ai	10			The title of the EVA has been changed without any communication to the stakeholders or those companies involved in the process. No reason has been given as to why. The original title was "EVA: Artificial Intelligence for analysing chest x-ray images to diagnose lung cancer". **This creates a number of fundamental issues with the EVA**. The original title and question created as a result have not been changed accordingly. The Decision questions and objectives as defined in 2.1 have not been changed and are therefore compromised and therefore invalid. The eligibility criteria in 3.1.2 are compromised and therefore invalid. This would lead the reviewers to conclude that there is no evidence available to determine a early value assessment of AI technologies used to diagnose lung cancer. Therefore the conclusion on page 9-10 cannot be accepted as correct. ***The review has been compromised***.	The title of the report is consistent with that of the final scope, as published on the NICE website
Behold.ai	11			We note that the NICE Specialist Committee members have advised Warwick Evidence. We have repeatedly expressed our concerns of the conflicts of interests of at least two members of this committee. ***Therefore the integrity of this EVA has been compromised.***	The EAG is not involved in identifying or recruiting members to the NICE specialist committee.
Behold.ai	12			We note the definition of artificial intelligence: "Artificial intelligence is a part of computer science where computers do some tasks that are usually done by humans." behold.ai are the only company registered with the Care Quality Commission to use its technology to provide a autonomous diagnosis to rule out normal chest x-rays using our CE Class IIa device on the GP, A&E and Outpatient pathway. This has been commercially rolled out for over 18 months.	The interventions included in the review are those specified in the NICE scope.

Artificial Intelligence software for analysing chest X-ray images to identify suspected lung cancer : early value assessment

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				Please find attached a prospective validation of the technology at Somerset and Taunton NHS Trust presented at Cancer Research UK Scientific Meeting in November 2022. This supports the position as the only company with technology that can be deemed as Artificial Intelligence. This has not been mentioned in this EVA and should be *immediately included* as to address our issues as highlighted in point 1.	The attached study was not identified by the ERG searches and was not provided in the company submission. However, it would be excluded on the interventions eligibility criteria as it does not compare AI+ reader versus reader alone
Behold.ai	13			The review includes "AI-Rad Companion Siemens.26 (Siemens 2022) It is unclear whether the prototype AI software described in Siemens 2022 is commercially available". It should therefore be removed from the EVA.	This study was provided in the Siemens company submission and was included for completeness with the appropriate caveat.
Behold.ai	14			Test accuracy results from summarised (but ineligible) studies section, Table 3, should include only papers that have used NHS data. Our issues highlighted in point 1 has resulted in three of our papers being excluded they should be immediately included to address our issues in point 1.	The *a priori* eligibility criteria defined in the review protocol did not specify use of NHS data. Post hoc eligibility criteria were developed in discussion with NICE to broaden the criteria and allow studies where the referral status of the population was unclear. This is a stated limitation of the review. The studies provided by the company were excluded by the protocol criteria, rather than having unclear eligibility.
Behold.ai	15			Section 4.6 "We identified one ongoing trial (KCT0005466) comparing Lunit INSIGHT in conjunction with a radiologist versus radiologist alone, however the population is those undergoing CXR for any reason in the outpatient department.". Therefore Lunit should be removed from the EVA.	Post hoc eligibility criteria were developed in discussion with NICE. This study was summarised with the appropriate caveat.
Behold.ai	16			Section 5.3. Clinical pathway for representation in model notes that the rule out or normal pathways. As per out points in 1, the eligibility criteria has been compromised and therefore three of our papers have been excluded. They should be *immediately included* as to address our issues as highlighted in point 1 and point 3.	Thank you for your feedback. The clinical pathways reflect the trajectory of patients after CXR review by radiologist/ reporting radiographer with or without AI software.
Behold.ai	17			There is no evidence eligible or ineligible supplied by Annalise.AI in the EVA and therefore this company should be removed	Annalise.AI was specified on the NICE scope and ERG protocol. No change to be made.

Artificial Intelligence software for analysing chest X-ray images to identify suspected lung cancer : early value assessment

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
Behold.ai	18			**Reference Attachments** 1. Cancer Research UK Lung Cancer Conference – Poster on Somerset NHS Trust: 2. Dyer T, Smith J, Dissez G, et al. Robustness of an Artificial Intelligence Solution for Diagnosis of Normal Chest X-Rays. Published online August 31, 2022. doi:10.48550/arXiv.2209.09204 3. Dyer T, Dillard L, Harrison M, et al. Diagnosis of normal chest radiographs using an autonomous deep-learning algorithm. *Clinical Radiology*. 2021;76(6):473.e9-473.e15. doi:10.1016/j.crad.2021.01.015 4. Dissez G, Tay N, Dyer T, et al. Enhancing Early Lung Cancer Detection on Chest Radiographs with AI-assistance: A Multi-Reader Study. Published online August 31, 2022. doi:10.48550/arXiv.2208.14742 5. Tam MDBS, Dyer T, Dissez G, et al. Augmenting lung cancer diagnosis on chest radiographs: positioning artificial intelligence to improve radiologist performance. *Clinical Radiology*. 2021;76(8):607-614. doi:10.1016/j.crad.2021.03.021	
The Society & College of Radiographers Dr [redacted]	19	6	Background para 2	Reporting radiographers as well as radiology specialists will use AI to help identify lung cancer on chest X-rays using AI	Thank you for this comment. The report should have said radiologist/reporting radiographer. We have corrected this
The Society & College of Radiographers Dr [redacted]	20	7	Methods, intervention/comparator	Reporting radiographers as well as radiology specialists will use AI to help identify lung cancer on chest X-rays using AI	Thank you for this comment. Please see response to comment 19.

Artificial Intelligence software for analysing chest X-ray images to identify suspected lung cancer : early value assessment

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
The Society & College of Radiographers Dr [redacted]	21	17	1.4	Reporting radiographers as well as radiology specialists will use AI to help identify lung cancer on chest X-rays using AI	Thank you for this comment. Please see response to comment 19.
The Society & College of Radiographers Dr [redacted]	22	19	1.7	High risk patients with a normal CXR may also be referred for a CT chest (with or without contrast) by primary care, there are increasing numbers of Trusts with this GP direct access [see NHS Digital DID]	Thank you for informing us about this.
The Society & College of Radiographers Dr [redacted]	23	22	Table 1	Reporting radiographer interprets chest X-rays (not radiographer)	Thank you for highlighting this error. We have corrected this.
The Society & College of Radiographers Dr [redacted]	24	25	3.2.1 Para 2	We have planned a health economic analysis as part of our recently approved (IRAS 317009; REC and HRA 21 Feb 2023) prospective study examining the impact of AI triage of GP chest X-rays on the time to diagnosis of lung cancer. The HEA includes a discrete choice experiment on patient/carer preferences.	Thank you for letting us know about this. We look forward to reading this work.
The Society & College of Radiographers Dr [redacted]	25	43	4.6	We have a recently approved (IRAS 317009; REC and HRA 21 Feb 2023) prospective study examining the impact of AI triage of GP chest X-rays on the time to diagnosis of lung cancer.	Thank you for letting us know about this. We look forward to reading this work.
The Society & College of Radiographers Dr [redacted]	26	58	5.4 Intermediate measure TAT	Times included for CXR review seem to be taken from experimental studies with a binary decision (cancer/no cancer, nodule/no nodule) whereas in routine clinical practice a report is required to address any and all findings including an answer to the clinical question posed in the referral. 20-30 seconds per chest X-ray report in clinical practice is likely to be unachievable The figure provided by the RCR [ref 41] has been misinterpreted – this is 80 reports PER SESSION (standard consultant PA DCC is 4 hours) not PER HOUR. Therefore 45 seconds per chest X-ray is incorrect and should be 3 MINUTES (80 reports in 240 minutes)	Thank you for this feedback and for highlighting our error. We have corrected this.

Artificial Intelligence software for analysing chest X-ray images to identify suspected lung cancer : early value assessment

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
The Society & College of Radiographers Dr [redacted]	27	60	5.4 Intermediate measure number of CR referrals	The number of patients referred for CT and the time from chest X-ray to CT is a secondary outcome of our prospective study	We look forward to reading this work.
The Society & College of Radiographers Dr [redacted]	28	62	5.4 Intermediate measure number of cancers missed	Patients with a chest X-ray included in our prospective study will have local hospital data checked for subsequent CT chest and help identify the number of missed lung cancers on GP referred chest X-rays	We look forward to reading this work.
The Society & College of Radiographers Dr [redacted]	29	63	5.4 Intermediate measure time to CT	Time from chest X-ray to CT chest is a co-primary outcome in our prospective trial that will involve 8 NHS sites in England	We look forward to reading this work.
The Society & College of Radiographers Dr [redacted]	30	64	5.4 Intermediate measure time to diagnosis	Time to diagnosis of lung cancer is a co-primary outcome in our prospective trial that will involve 8 NHS sites in England	We look forward to reading this work.
The Society & College of Radiographers Dr [redacted]	31	66	5.4 Summary	We have recently obtained REC and HRA approval for a prospective, block randomised trial at 8 NHS sites that will determine the clinical (time to diagnosis of lung cancer, time from chest X-ray to CT chest) and cost-effectiveness (HEA and DCE) of AI triage of GP chest X-rays for immediate report. This will run for 12 months, with an estimated 150,000 CXRs included.	We look forward to reading this work.
The Society & College of Radiographers Dr [redacted]	32	72	Q5 costs & resource of using AI	We are conducting an Evaluation of staff acceptability, experience and confidence in deployment of QXR AI triage tool as part of our prospective clinical trial	We look forward to reading this work.
The Society & College of Radiographers Dr [redacted]	33	76	Q5 costs of further tests	Our recently approved prospective multisite study includes number of CR scans performed, number of normal/discharged patients, number of cancers missed and detected and stage at diagnosis of lung cancer	We look forward to reading this work.

Artificial Intelligence software for analysing chest X-ray images to identify suspected lung cancer : early value assessment

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
The Society & College of Radiographers Dr [redacted]	34	78	Table 10	Reporting radiographer costs are likely to be significantly lower than consultant radiologist (AfC Band 7-8b, most B7)	Thank you for providing this information. We provided indicative costs for consideration in any future modelling.
The Society & College of Radiographers Dr [redacted]	35	86	6.1.2	We have recently obtained REC and HRA approval for a prospective, block randomised trial at 8 NHS sites that will determine the clinical (time to diagnosis of lung cancer, time from chest X-ray to CT chest) and cost-effectiveness (HEA and DCE) of AI triage of GP chest X-rays for immediate report. This will run for 12 months, with an estimated 150,000 CXRs included.	We look forward to reading this work.
The Society & College of Radiographers Dr [redacted]	36	89	6.3	[comment incomplete - to include Qure papers?]	
The Society & College of Radiographers Dr [redacted]	37	89	6.4	EDI – this should include known bias in lung cancer and artificial intelligence not just the research team. Significant inequality in outcomes for lung cancer patients are found in the UK with lung cancer having one of the greatest deprivation gaps compared with other cancers [UK Lung Cancer Coalition, 2022]. Lung cancer disproportionality affects lower socioeconomic populations, regional location, certain ethnic groups and those that may not frequently engage with health services. Although multifactorial, access to screening and diagnostic services are one element driving health inequalities in lung cancer outcomes, and making the pathway as short as possible will play a role. A single diagnostic episode, with patients with a CXR suspicious for lung cancer offered a same day CT scan, is one way to increase engagement for populations who may not routinely engage with health services and removes communication barriers that may prevent follow up attendance.	Thank you for highlighting these very relevant points.

Artificial Intelligence software for analysing chest X-ray images to identify suspected lung cancer : early value assessment

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				Significant bias has been identified in training and test sets used to develop and validate AI chest X-ray algorithms. Gichoya et al (2022) found AI identified race in medical imaging which is “an enormous risk for all model deployments”, Larrazabal et al (2020) report significant gender imbalance. If we are to improve health outcomes for all lung cancer patients we need to ensure that we reduce not entrench health inequalities.	
The Society & College of Radiographers Dr [redacted]	38	91	7.2	Our recently approved prospective, multisite block randomised trial will address the following research priorities: Assessment of test accuracy and test failure rates Assessment of AI on clinical decision-making and acceptability Time from chest X-ray to diagnosis of lung cancer with intermediate outcomes (time to CT, number of respiratory referrals) across the full pathway Discrete choice experiment on patient/carer perspectives on AI for lung cancer diagnosis on chest X-rays	We look forward to reading this work.
British Thoracic Society	39			“The EVA process is helpful for looking at new and emerging technologies. This report highlights limitations with the applicability of the AI software in analysis of CXR images to NHS practice at this point in time. Some comments about the report:	Thank you for your feedback. We agree with all of your comments.
British Thoracic Society	40			No study included more than 1 software- it would be good to see head to head comparisons to know how the different software may perform in the same scenarios.	
British Thoracic Society	41			No data were available on clinical outcomes or technical failure rates, which is important if a system were to be deployed in the NHS.	
British Thoracic Society	42			Only 1 study was in a UK population so applicability (in a population with lower prevalence of granulomatous disease) would be important to understand.	
British Thoracic Society	43			The cost effectiveness assessments are based on only 1 NHS Trust which makes it difficult to be certain how robust this assessment is.	

Artificial Intelligence software for analysing chest X-ray images to identify suspected lung cancer : early value assessment

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
British Thoracic Society	44			It seems clear that further research is needed to assess the role of these AI solutions in the UK population and NHS based on this report.”	