DIAGNOSTICS ASSESSMENT PROGRAMME

Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals

Early value guidance consultation document – Comments

THEME: RECOMMENDATIONS

Comment number	Name and organisation	Section number	Comment	NICE response
1	The Royal College of Radiologists	1.1	Yes, we think this is a reasonable recommendation, whilst prospective evidence is gathered as research/pilot study. However, it will be important to ensure that the evidence is gathered in a systematic way to ensure that the resulting data are correct / fair/ unselected / unbiased. The data collected should be consecutive and the patient follow up time should be sufficient to ensure that misses are identified through subsequent patient follow-up. In addition, false positives must also be carefully collected. Planning the research properly is really important.	Thank you for your comment which NICE has considered. The early value guidance specifies the need for prospective studies (sections 3.7 and 3.11).
2	The Royal College of Radiologists	1.2	Yes, we agree.	Thank you for your comment which NICE has considered.
3	The Royal College of Radiologists	1.3	Yes, we agree.	Thank you for your comment which NICE has considered.

DIAGNOSTICS ASSESSMENT PROGRAMME

Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals

Early value guidance consultation document – Comments

THEME: EVIDENCE ASSESSMENT

Comment number	Name and organisation	Section number	Comment	NICE response
4	The Royal College of Radiologists		We are not aware of any other large studies specifically on lung cancer AI in CXR.	Thank you for your comment which NICE has considered.
5	The Royal College of Radiologists		The reference standard for the final correct diagnosis (lung cancer or no lung cancer) is not defined and this is essential for a diagnostic test accuracy study. Defining the reference standard is missing in this document, particularly as there is discussion of the potential changes in sensitivity and specificity for detection – possible increase in unnecessary CT scans etc. The final reference standard is needed to calculate the difference in sensitivity and specificity, and accuracy between the intervention and the comparator. This is essential to answer point 3.13, the clinical effectiveness. Perhaps this is meant to be included as part of any research protocol, but it seems that this should be included alongside the description of the interventions and the comparator. Separate from the guidance consultation document, a reference standard appears to be suggested by Table 1 of the diagnostics assessment report.	Thank you for your comment which NICE have considered. We agree that the reference standard for any accuracy studies on AI-derived software for analysing chest X-rays needs to be in study protocols. The EAG included studies in the assessment report that used the following reference standards: • For accuracy of lung cancer detection: Lung cancer confirmed by histological analysis of lung biopsy, or diagnostic methods specified in NICE guideline 122, where biopsy is not applicable. • For accuracy of nodule detection: Radiology specialist (single reader or consensus of more than one reader).

DIAGNOSTICS ASSESSMENT PROGRAMME

Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals

Early value guidance consultation document – Comments

THEME: EVIDENCE ASSESSMENT

Comment number	Name and organisation	Section number	Comment	NICE response
6	Lunit Ltd	External Assessment Report. Key issue 1: Lack of evidence about diagnostic accuracy, technical failure rates, clinical decision making and clinical outcomes	We would like to draw attention to a couple of evidence omitted by the assessment team. The draft document mentioned that: • There is a lack of evidence on diagnostic accuracy, clinical decision making and clinical outcomes The Peer Reviewed Publication described below addresses this gap: Lunit INSIGHT CXR Peer Reviewed Publication relating to **Incidental lung cancer detection** See publication attached as **evidence A**, also referenced below. **Publication URL**: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0281690 **DOI** https://doi.org/10.1371/journal.pone.0281690 ○ This publication specifically looks at the increase in the incidentally detected lung cancers after INSIGHT CXR adoption. You can see that the majority of incidentally detected lung cancers were early stage, showing the promise for the use case of CXR AI. ○ This is meaningful as this is a study at a site that is using INSIGHT CXR prospectively (but retrospectively analysed).	Thank you for your comment which NICE have considered. The EAG have noted that his paper was published after the searches for the review were conducted (March 2023 vs November 2022). The EAG also note that this study does not meet the review a priori or post hoc eligibility criteria, because it has no eligible comparator. Therefore, it would not have been included even if published earlier. The additional review by the EAG also excluded the study at full-text review as it was AI alone with no eligible comparator.

DIAGNOSTICS ASSESSMENT PROGRAMME

Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals

Early value guidance consultation document – Comments

THEME: EVIDENCE ASSESSMENT

Comment number	Name and organisation	Section number	Comment	NICE response
			We implore the committee to consider this evidence and revise their conclusions	
7	Lunit Ltd	External Assessment Report Key issue 2: Lack of evidence about time to read and report and acceptability to clinicians	We would like to draw attention to a couple of evidence omitted by the assessment team. The draft consultation document mentioned that: • There is a lack of evidence about time to read and report. There are a number of Peer Reviewed Publications on Lunit INSIGHT CXR that addresses these gaps as summarised below **Time to report reduction:** • This reader study looked at four findings, including nodules, showing that there was a significant time to report reduction with AI. See publication attached as **evidence C,** also referenced below. **Publication url:** https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2795798 **doi:**10.1001/jamanetworkopen.2022.29289 We implore the committee to consider this evidence and revise their conclusions	Thank you for your comment which NICE have considered. The EAG have noted that this study does not meet the review inclusion criteria. It is listed in appendix 2: Table of studies excluded at full text assessment in the assessment report. The reason for exclusion was that the chest X-rays were from two hospital databases (one is an intensive care database) and there were no details of the referral route of participants. During the addendum review, the study was excluded at title and abstract by the EAG as the comparison is clinician with AI-derived software versus clinician alone and therefore out of scope.

DIAGNOSTICS ASSESSMENT PROGRAMME

Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals

Early value guidance consultation document – Comments

THEME: EVIDENCE ASSESSMENT

Comment number	Name and organisation	Section number	Comment	NICE response
8	Lunit Ltd	External Assessment Report	We would like to draw attention to a couple of evidence omitted by the assessment team. The draft document speak about the: • Lack of prospective/RCT evidence The INSIGHT CXR Peer Reviewed Publication described below addresses this gap: **Prospective RCT: AI Improves Nodule Detection on Chest Radiographs in a Health Screening Population** ○ This study showed that the AI improved actionable and malignant nodule detection, without increasing the false referral rates, in a prospective RCT. **Full publication here**: See publication attached as **evidence B**, also referenced below. **DOI:** https://pubs.rsna.org/doi/10.1148/radiol.221894 We implore the committee to consider this evidence and revise their conclusions	Thank you for your comment which NICE have considered. The EAG have noted that this paper was published after the searches for this review were conducted (February 2023 vs November 2022). The EAG also note that this study comprised a health screening population. This is not within the scope of the assessment, so would have been excluded even if published earlier.
9	Healthcare professional	External Assessment Report Table 4	As requested by NICE team via email we submitted our recently approved study protocol and HRA/REC approvals for LungIMPACT. Our work will answer many of the areas identified as in need of further evidence. LungIMPACT has not been included in the summary of ongoing studies. ISRCTN registration has been submitted	Thank you for your comment which NICE have considered. This information has been added to the guidance document in section 3.11: An ongoing randomised controlled trial

DIAGNOSTICS ASSESSMENT PROGRAMME

Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals

Early value guidance consultation document – Comments

THEME: EVIDENCE ASSESSMENT

Comment number	Name and organisation	Section number	Comment	NICE response
			The trial registration for LungIMAPCT came through this morning if helpful https://www.isrctn.com/ISRCTN78987039	(LungIMPACT) looking at the impact of AI-derived software for triaging to chest CT in people who have been referred by their GP for chest X-rays may provide useful evidence. Key outcomes are time to CT scan and time to lung cancer diagnosis.
10	Behold.ai	3.1.2	Eligibility Criteria - Table 1 defines three key questions relating to AI software. Exclusion criteria defined areas that NICE deem inclusive are within "*NICE real-world evidence framework"* published 23rd June 2022. It is clear this conflicts with the papers research approach and eligibility.	Thank you for your comment which NICE have considered.
11	Behold.ai	External Assessment report	"Titles and abstracts of records identified by the searches were screened by one reviewer." We feel this approach allowed for studies from Behold.ai to be deliberately excluded from the review. "Risk of bias" was assessed by one reviewer. Again we feel this approach allowed for studies from Behold.ai to be deliberately excluded from the review	Thank you for your comment which NICE have considered. Early value assessments are not intended to replicate the methods of a systematic review. This is because of the short for the assessment. Further details of the EVA process in NICE's early value assessment interim statement. The limitations, and potential impacts, of the review methods are described in the final report discussion.

DIAGNOSTICS ASSESSMENT PROGRAMME

Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals

Early value guidance consultation document – Comments

THEME: EVIDENCE ASSESSMENT

Comment number	Name and organisation	Section number	Comment	NICE response
12	Behold.ai		Exhibit 6,7,8,9,10 – Behold.ai studies presented to the Committee but not included. Our studies demonstrate the only real-world evidence from an NHS setting that has been presented to the committee, yet this not been included in the review process and once again is a deliberate action to exclude significant and relevant data from the review. Exhibits 5 and 10 highlight clinical effectiveness to deliver rule-out normal diagnosis. Our paper has not been included within the review of literature, our paper demonstrates multiple system and patient benefits.	Thank you for your comment which NICE have considered. The EAG noted that the exhibits mentioned were excluded for the following reasons: Exhibit 6: A poster not a peer reviewed publication; does not compare radiology specialist + AI to radiology specialist alone. Exhibit 7: Does not compare radiology specialist + AI to radiology specialist alone; chest X-rays were chosen to represent a diverse dataset of NHS patients and care settings. Exhibit 8: Population includes A&E, GP, outpatient. Exhibit 10: Population referral route not reported. Includes chest X-rays with difficult to locate nodules and chest X-rays with no nodules.

DIAGNOSTICS ASSESSMENT PROGRAMME

Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals

Early value guidance consultation document – Comments

THEME: EVIDENCE ASSESSMENT

Comment number	Name and organisation	Section number	Comment	NICE response
				Exhibit 9 was included in the original post-hoc analysis. Exhibit 10 was included in the addendum analysis.
13	Behold.ai		Multiple studies were submitted for review and Table 2 and Appendix 2 details exclusion reasons. We refer to Exhibits 6,7,8,9,10 that refer to abstract from real-world study at Somerset NHS Trust. Real-world evidence from NHS sites are only available from Behold.ai vs the competitors. The real-world evidence is aligned to "*NICE real-world evidence framework*". The exclusion of our rule normal evidence includes two peer reviewed and published papers which we can confirm includes NHS data on the GP pathway (please see Exhibit 6,7,8,9,10) and real-world evidence from Somerset and Taunton NHS hospital presented at Cancer Research UK Scientific Meeting in November 2022 (Exhibit 6). The reason for the exclusion is "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." This does not make any logical sense noting the title of the EVA is "artificial intelligence-derived software to analyse chest x-rays for	Thank you for your comment which NICE have considered. The EAG noted that the exhibits mentioned were excluded for the following reasons: Exhibit 6: A poster not a peer reviewed publication; does not compare radiology specialist + AI to radiology specialist alone Exhibit 7: Does not compare radiology specialist + AI to radiology specialist alone; chest X-rays were chosen to represent a diverse dataset of NHS patients and care settings. Exhibit 8: Population includes A&E, GP, outpatient. Exhibit 10: Population referral route not reported. Includes chest X-rays with

DIAGNOSTICS ASSESSMENT PROGRAMME

Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals

Early value guidance consultation document – Comments

THEME: EVIDENCE ASSESSMENT

Comment number	Name and organisation	Section number	Comment	NICE response
			suspected lung cancer in primary care referrals." **Not** its original title ""Artificial Intelligence for analysing chest x-ray images to **diagnose** lung cancer".	difficult to locate nodules and chest X-rays with no nodules. Exhibit 9 was included in the original post-hoc analysis. Exhibit 10 was included in the addendum analysis.
14	Behold.ai		Behold.ai have an approved business case identifying cost effectiveness to the NHS Trusts. Data on cost-effectiveness studies was not requested by NICE.	Thank you for your comment which NICE have considered.

DIAGNOSTICS ASSESSMENT PROGRAMME

Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals

Early value guidance consultation document – Comments

THEME: EQUALITY ISSUES

Comment number	Name and organisation	Section number	Comment	NICE response
15	The Royal College of Radiologists		"NICE: Equality issues? No issues that we can see."	Thank you for your comment which NICE has considered.

DIAGNOSTICS ASSESSMENT PROGRAMME

Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals

Early value guidance consultation document – Comments

THEME: GENERAL ISSUES

Comment number	Name and organisation	Section number	Comment	NICE response
16	Healthcare professional	Table of contents	Page numbers for many sections are incorrect with duplicate numerals	Thank you for your comment which NICE has considered.
17	Behold.ai	NICE process	Composition of Selection Committee with conflict of interests: behold.ai technologies limited (behold.ai) started to raise concerns with the integrity of the process from the 25th November 2022 with regards to the inclusion of Professor Baldwin as a member of the Specialist Committee. This was based on his working relationship with qure.ai, a company that claims to have a similar product to behold.ai (**please see Exhibit 1 – Conflicts of Interest – Specialist Committee – Baldwin and Exhibit 2 - Prof David Baldwin** - SBRI Healthcare announces £3.2 million in Funding for Qure.ai's A.pdf). As a result of our highlighting this clear conflict of interest, we received confirmation on 25th January that Professor Baldwin had voluntarily stood down from the Specialist Committee. It was disappointing and surprising that this clear conflict of interest had not been identified beforehand and that it took almost three months for this to be acknowledged and resolved. Further, we also raised legitimate concerns about the position of Dr Gleeson on the Special Committee on 25th January – please refer to **Exhibit 3**. These concerns were rebuffed and not satisfactorily addressed.	Thank you for your comment which NICE have considered. Professor Baldwin voluntarily stood down from the committee and Professor Gleeson attended part 1 of the committee meeting to give expert advice but was not a decision maker and did not attend part 2 of the committee meeting. The Royal Collage of Radiologists were invited to participate in this assessment as a stakeholder.

DIAGNOSTICS ASSESSMENT PROGRAMME

Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals

Early value guidance consultation document – Comments

THEME: GENERAL ISSUES

Comment number	Name and organisation	Section number	Comment	NICE response
			We and our advisers find it astonishing that the National Institute for Health and Care Excellence (NICE) reasonably thought that these two prominent members of the Specialist Committee were ever considered to be appropriate to undertake an independent valuation of these technologies given their clear conflicts of interest. Our concerns over the integrity of this EVA were further heightened when we noticed that the Royal College of Radiologists seemed to have been excluded from the EVA process (**Please see Exhibit 4** – Communication with RCR President re Stakeholder participation).	
18	Behold.ai	Scope	Decision and Objectives – change in scope. We noted, only via public disclosure on the NICE website, that the title of the EVA had changed from "Artificial Intelligence for analysing chest x-ray images to diagnose lung cancer" to "using artificial intelligence-derived software to analyse chest x-rays for suspected lung cancer in primary care referrals" without any consultation with the stakeholders including behold.ai. We escalated this issue with the fundamental change of the evaluation scope, noting it had now included a specific pathway in its title, namely the inclusion of primary care referrals.	Thank you for your comment which NICE have considered. The final scope was published on the NICE website on 22 November 2022 with the title: Artificial Intelligence software for analysing chest X-ray images to identify suspected lung cancer. Registered stakeholders were informed by email about the publication of the scope. The change to the title between the draft scope and the final scope was based on discussions with clinical experts at the

DIAGNOSTICS ASSESSMENT PROGRAMME

Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals

Early value guidance consultation document – Comments

THEME: GENERAL ISSUES

Comment number	Name and organisation	Section number	Comment	NICE response
			We never received an explanation until the publication of Diagnostic Consultation on the 21st March 2023 which stated "The interventions included in the review are those specified in the NICE scope." (https://www.nice.org.uk/guidance/indevelopment/gid-dg10065/documents) It clearly is not the case, that the change of scope and title means that the interventions included are in scope as a result. An exclusion criteria has been created with the change of this title without a process of reassessment of the clinical literature available and inclusion and exclusion criteria.	assessment subgroup meeting held on 15 November 2022.
19	Behold.ai	External Assessment Report and Scope	Decision and Objectives – change in scope as stated in "General (2)". The initial request from suppliers had the provisional title "*Artificial intelligence for analysing chest x-ray images to diagnose lung cancer (provisional title)*". No information was shared with suppliers to state the title, decision objectives and questions had been changed. This impacts the core concepts of the paper throughout the analysis and conclusions.	Thank you for your comment which NICE have considered. The final scope was published on the NICE website on 22 November 2022 with the title: Artificial Intelligence software for analysing chest X-ray images to identify suspected lung cancer. The change to the title between the draft scope and the final scope was based on discussions with clinical experts at the assessment

DIAGNOSTICS ASSESSMENT PROGRAMME

Artificial intelligence-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals

Early value guidance consultation document – Comments

THEME: GENERAL ISSUES

Comment number	Name and organisation	Section number	Comment	NICE response
				subgroup meeting held on 15 November 2022.
20	Behold.ai	External Assessment Group	On 3rd March 2023 a Committee Meeting was held, which behold.ai attended. Behold.ai raised several issues but most importantly, it was clear that our technology had not been correctly reviewed by the EAG. Please see '**Exhibit 5** – CADX rule out normal exclusion of evidence question 1' which clearly states that we told the EAG that our medical device has the capability ***to rule out cancer using our high confidence autonomous findings with a fully authorised and transcribed radiology report, a CADX device.*** We were given the opportunity to highlight this factual inaccuracy at the end of the open session, however on reviewing the diagnostic assessment report dated the 21st March 2023, this factual inaccuracy has not been addressed. (**Please see Exhibit 11**)	Thank you for your comment which NICE have considered. Thank you for highlighting this factual inaccuracy which will be updated in the final guidance.
21	Behold.ai	General	This catalogue of fundamental errors leads us to conclude that the exclusion of our evidence which would support the recommendation of behold.ai technology, as part of this EVA, **is deliberate**.	Thank you for your comment which NICE have considered.