Newcastle University

NHS
The Newcastle upon Tyne Hospitals
NHS Foundation Trust

GID- HTE10059 Artificial intelligence technologies to aid the opportunistic detection of vertebral fragility fractures (VFF) on radiographic images

EAG comments on additional evidence by Nanox AI

Produced by	Newcastle External Assessment Group (EAG)
Main Authors (Name, position)	**Paula Leslie**, Pre-registrant Clinical Scientist, NuTH **Rosalyn Parker,** Evaluation Healthcare Scientist, NuTH **Kim Keltie,** Director, NuTH
Correspondence to	nuth.nmpce.hta@nhs.net
Completion date	14/08/2025

Declared competing interests of the authors: None

Responsibility for report:

The views expressed in this report are those of the authors and not those of NICE. Any errors are the responsibility of the authors.

None of the information included in this report is considered academic in confidence or commercial in confidence.

1. Aim

Following the stakeholder consultation for the early value assessment (EVA) for artificial intelligence (AI) technologies to aid opportunistic detection of vertebral fragility fractures (VFF) on radiographic images, two additional sources of evidence were submitted by Nanox AI for (HealthOST) to be considered by the Committee (Mckeag et al. 2024, Mathew et al. 2025). The External Assessment Group (EAG) were subsequently asked by NICE to provide a high-level narrative summary of the evidence and quality of evidence. This addendum provides a high level summary of the key findings and limitations of the two studies provided.

2. Mathew et al. 2025

The EAG note that the publication by Mathew et al. (2025) was published in June 2025, which was after the EAG searches conducted in February 2025 (see Appendix A1 of the EAG Report). The paper reports on the detection of incidental vertebral compression fractures in a retrospective analysis of 675 outpatient cases with chest, abdomen or pelvis CT scans (conducted between February 2019 and March 2020) in people aged 50 years or over (which the EAG note is in line with the technologies instructions for use), from a single hospital in Canada. HealthOST (Nanox AI) was applied to 610 consecutive cases after 65 cases were excluded due to non-compliance with the algorithm requirements; with reported reasons including: less than 15 cm of the spine was detected (34 cases); fewer than four vertebrae were observed (13 cases); there was an absence of a valid CT series (7 cases); there was an insufficient number of images, specifically less than 20 (8 cases); and systemic error (3 cases). A further 20 cases were excluded due to subsequent identification of bone metastases or spinal hardware. The EAG note that the reasons provided are a mixture of appropriate exclusions based on the contraindications listed in the IFU, technical failures meaning the technology was unable to process the image and additional clinical reasons; therefore the technical failure rate is unclear.

Two radiologists (with specialist musculoskeletal experience) reviewed images and reached consensus with discussion (reference standard) blinded to the HealthOST rating. A third specialist was consulted if consensus could not be reached. AI findings were compared to the original clinical reports. Analysis was based on

individual vertebrae rather than patient cases, and at two thresholds for AI fracture detection. For 20% AI fracture detection: sensitivity was reported at 92.0% and specificity at 52.7%. For 25% AI fracture detection: sensitivity was reported at 78.0% and specificity at 94.2%. The false positive rates were 42.9% at 20% threshold and 5.4% at 25% threshold.

This published paper provides additional evidence for HealthOST in an international outpatient setting and comparison with standard care. Authors acknowledge reliance on intravertebral evaluation as a limitation (rather than comparing to adjacent vertebrae), which they state could lead to discrepancies in fracture grading and overcalls. Authors state that HealthOST required at least 4 contiguous vertebrae for accurate segmentation, stating that Nanox AI was working on adjusting software to address edge of field of view overcalls.

3. Mckeag et al. 2024

Mckeag et al. (2024) is a non-peer reviewed poster available on the Company website and was not submitted to the EAG for consideration during the assessment or stakeholder consultation period. The poster has limited detail of the methods and no information on: reason for imaging, patient characteristics, dates of imaging, original radiology report. The poster presents diagnostic data presumed to be retrospective and from a UK NHS setting given author affiliation. HealthOST (Nanox AI) was applied to 619 CT images which included T1-L4 vertebrae (anatomical location of included scans not reported) of which 597 were successfully analysed for existence and severity of vertebral compression fractures (VCF). This represents a failure to anlaysis of 3.6% (22/619). The HealthOST outputs were compared with reports from an independent consultant radiologist reporting VCFs, blinded to the HealthOST outcome, with 95.6% (571/597) having full agreement. There was agreement on presence of fracture but disagreement on severity in 11 additional images, and the EAG assumes disagreement on the remaining 15 images. A random 10% sample plus the images lacking agreement were reviewed by two trainees (profession and degree of musculoskeletal training not reported) who were blinded to prior reports. The trainees agreed with HealthOST in 10 and 12 discordant

cases (unclear reporting). Sensitivity of HealthOST was reported at 81.4% and specificity at 99.6%, however it is unclear from reporting how this was calculated and what was considered the reference standard. The poster does not report on the agreement between the trainees and the original consultant radiology report. This poster provides additional evidence for HealthOST in an NHS setting but is lacking peer review (available only on the company website) and detail.

4. References

Mathew V, Pearce D, Kates Rose N et al. (2025) Clinical validation of commercial AI software for the detection of incidental vertebral compression fractures in CT scans of the chest and abdomen. Diagnostics (Basel) 15(12):1530

Mckeag CP, Friar M, Baker P, et al. (2024) Validation of HealthOST by Nanox.AI – image processing software that provides qualitative and quantitative analysis of the spine from CT images to support clinicians in the evaluation and assessment of musculoskeletal disease of the spine [Online; accessed 13 August 2025]