Synthetic Demonstration Manuscript (SAFE TO SHARE)

PET/CT AI Triage and Resident Education: A Simulated Multi-Site Reader Study

NOTE: This document is entirely fictional and contains no real data, patients, or institutions.

Date: February 05, 2026

Abstract

Background: Demonstrate the local reviewer pipeline on synthetic content. Purpose: Evaluate a simulated AI-assisted PET/CT triage workflow and a resident education module. Materials and Methods: We created a fictional retrospective multi-site dataset (N=420 patients) with three scanners, two radiopharmaceuticals, and a reader study with 6 readers (3 faculty, 3 residents). Data were split at the patient level into train/validation/test with an external validation site. Primary endpoint was diagnostic accuracy for lesion-level detection; secondary endpoints included calibration and time-to-report. For education, we assessed pre/post knowledge with a 20-item test and collected satisfaction. Results: The AI model achieved AUROC 0.91 in internal testing and 0.87 at the external site. Calibration slope was 0.92. Median interpretation time decreased by 14%. Residents improved by 12 percentage points on the knowledge test. Conclusion: In this synthetic example, AI-assisted triage and a short curriculum may improve efficiency and educational outcomes; real-world validation is required.

Introduction

Clinical adoption of AI in nuclear medicine imaging requires clear evidence of benefit, generalizability, and workflow feasibility. Education research similarly requires rigorous design and validity evidence. This synthetic manuscript exists solely to test the reviewer tool’s two-pass outputs and reporting guideline checks.

Methods

Study design and population: Fictional retrospective multi-site cohort with inclusion/exclusion described only partially for demonstration. Imaging protocol: PET/CT acquisition parameters are intentionally incomplete to trigger nomenclature and protocol comments. Radiopharmaceutical naming is intentionally inconsistent (e.g., “FDG tracer” and “Ga-PSMA”) to test the nomenclature module. AI model: Convolutional neural network with patient-level split; a potential leakage risk is introduced via repeated studies per patient and unclear handling. Statistics: Confidence intervals are reported for AUROC but not for sensitivity/specificity. Missing data handling is not specified. Education study: Pre/post design without a control group; assessment validity evidence is not described. No sample size justification is provided.

Results

Primary performance: AUROC values are reported; sensitivity/specificity at a decision threshold are shown in Table 2 (threshold selection unclear). External validation is described but details of site differences are sparse. Subgroup performance by scanner is partially described without multiple comparisons control. Education outcomes: Mean pre/post scores are reported without distributional checks; effect sizes are not reported.

Discussion

The discussion emphasizes potential workflow benefit and education improvement but under-addresses limitations, including selection bias, protocol variability, generalizability, and assessment validity. The manuscript includes claims of novelty but provides incomplete positioning against prior work.

Tables

Table 1: Cohort characteristics (synthetic). Table 2: Performance metrics by site. Table 3: Education outcomes.

Figures

Figure 1: Flow diagram of cohort selection (simplified). Figure 2: ROC curve and calibration plot (synthetic).

References

References are placeholders and include some intentionally incomplete citations to trigger reference-hygiene flags.