

# The Radiographer Image Interpretation Training Gap: A Systematic Review of Diagnostic Competency with Global Comparison and Implications for Workforce Extension in Nigeria

Agwu John Emeka<sup>1</sup>, Chigbo Okorie Kalu<sup>2</sup>, Okike Ikenna<sup>3</sup>, Yetunde Roselyn Ihimekpen<sup>4</sup>

<sup>1</sup>Department of Radiology, Obafemi Awolowo University Teaching Hospital Ife Nigeria.

<sup>2</sup> Nigerian Navy Hospital Borokiri

<sup>3</sup>Sheffield Teaching Hospital NHS Foundation Trust England

<sup>4</sup>Royal Bolton Hospital NHS Foundation Trust

Received: 11.09.2026 | Accepted: 09.10.2026 | Published: 15.10.2026

\*Corresponding author: Agwu John Emeka

DOI: [10.5281/zenodo.23386166](https://doi.org/10.5281/zenodo.23386166)

## Abstract

## Original Research Article

**Introduction:** Nigeria has fewer than 700 radiologists for more than 200 million people. The shortage is about 30% and may reach 40% by 2028 (Royal College of Radiologists, 2023). X-ray reports are significantly delayed. Nigeria has about 5,000 radiographers who take most X-rays, but they do not routinely report them. Radiographers are trained, in some countries to read and report images. We wanted to learn from global experience and see what could work in Nigeria. Methods: We followed PRISMA 2020, SWiM and registered our protocol PROSPERO CRD420261456144. Five databases (PubMed, Scopus, Embase, Web of Science, CINAHL) plus reference lists and citation tracking were searched. We found 1,207 records and included 23 studies after screening of the studies. We used TIDieR to describe training. Quality of included studies was checked using RoB2, ROBINS-I and QUADAS-2. Narrative synthesis (SWiM) was used to synthesize result because the tests were too different for meta-analysis. We used GRADE to rate certainty.

**Results:** Baseline skill without training was moderate. We found that 19 training studies showed improvement in the same direction and 15 were statistically significant. Risk of bias was 7 Low, 10 Moderate/Some concerns, 4 Serious and 2 High.

**Conclusion:** Baseline competency is moderate but short low-cost training improves it significantly and ongoing audit helps to keep skills. The Zambian WhatsApp model with data allowances appears most suitable for Nigerian district hospitals.

**Implications for practice:** We propose a phased model for Nigeria which include Phase 1, Phase 2, and Phase 3 which integrate Redot, Preliminary reporting and formal reporting practices. Regulation and budget for data and supervision are needed with regular audit for improvement.

**Keywords:** radiographer reporting; image interpretation; preliminary image evaluation; workforce extension; Nigeria; systematic review.

Copyright © 2026 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).



**Citation:** Agwu, J. E., Kalu, C. O., Okike, I., & Ihimekpen, Y. R. (2026). The radiographer image interpretation training gap: A systematic review of diagnostic competency with global comparison and implications for workforce extension in Nigeria. *SSR Journal of Medical Sciences (SSRJMS)*, 3(10), 43-66.

## Introduction

Medical imaging is important for modern healthcare. Plain X-ray is still the most common imaging test because it is available, quick and cheap (World Health Organization, 2023). The number of imaging requests is growing as populations increase and live longer, so radiology departments are under pressure. There are not enough radiologists worldwide, especially in low- and middle-income countries (Brady, 2017; Royal College of Radiologists, 2023).

Quick and correct reading of X-rays is very important in emergency care. Missed fractures are reported in 3-10% of emergency X-rays (Hallas and Ellingsen, 2006) and there is about 3% difference between emergency doctors and radiologists (Brunswick et al., 2004). In many healthcare facilities, particularly in resource-limited settings where radiologists are unavailable, clinicians are often required to interpret radiographs despite having limited formal training in image interpretation. These challenges have strengthened the argument for expanding the role of radiographers, who are generally more numerous than radiologists and are directly involved in image acquisition (Berman et al., 1985; Swinburne, 1971).

Traditionally only radiologists reported images. Because of high workload and staff shortages, radiographers have taken on more responsibility. The role started with the Red Dot system where radiographers marked abnormal X-rays (Berman et al., 1985). It then grew to preliminary image evaluation (PIE), where radiographers write a short description, and then to formal reporting where trained radiographers write the final report (Snaith and Hardy, 2008). These roles need good anatomy, pathology knowledge, visual search and reporting skills (Krupinski, 2010; Society and College of Radiographers, 2022).

Basic radiography training includes anatomy and imaging techniques, but training in image interpretation is very different in different universities and countries (Hardy and Snaith, 2009). Many training programmes have been developed including short courses, workplace

mentoring, online learning and postgraduate certificates (Hardy et al., 2013). Studies have shown these programmes improve accuracy, confidence and performance (Hargreaves and Mackay, 2003; Hazell et al., 2015; Neep et al., 2019).

However, programmes vary a lot in length, content, teaching method and assessment. Studies also measure success in different ways including accuracy, sensitivity, specificity and agreement with radiologists. This makes it hard to know which training works best. This is especially important for Nigeria. Nigeria has fewer than 700 radiologists for more than 200 million people and about 5,000 radiographers who do most X-ray work (Okeji et al., 2023; Radiographers Registration Board of Nigeria, 2022). Formal reporting by radiographers is not yet routine. Learning from countries where it works well could help with curriculum and workforce planning in Nigeria.

Previous reviews have looked at only one country or one aspect of reporting (Brealey et al., 2005; Murphy et al., 2019). They did not look at many types of training or at what could work in low-resource countries. The aim of this systematic review is to bring together global evidence on radiographer image interpretation training, check how it improves diagnostic skill, compare findings across countries, and possible applications for workforce extension in Nigeria.

## Methods

This review followed PRISMA 2020 guidelines (Page et al., 2021), the Synthesis Without Meta-analysis (SWiM) guide (Campbell et al., 2020) and our registered protocol PROSPERO CRD420261456144.

Eligibility criteria (PICOS): Population: Qualified diagnostic radiographers. We excluded radiologists, residents and other health workers unless radiographer data could be separated. Intervention: Any formal training to improve image interpretation including Preliminary image evaluation (PIE), Preliminary clinical evaluation (PCE), Red Dot and reporting courses. Comparator: Before and after training,

trained versus untrained, different types of training, or comparison with radiologist standard. Studies with no comparator were also included if they reported competency after training. Outcomes: Main outcomes were diagnostic accuracy, sensitivity, specificity and competency. Secondary outcomes were agreement with radiologists, localisation, description, confidence and error patterns. Study design: RCTs, quasi-experimental pre-post studies, cohort, diagnostic accuracy and audit studies. We excluded editorials, conference abstracts without full text, reviews, and non-English papers.

Information sources and search: We searched PubMed/MEDLINE, Scopus, Embase, Web of Science Core Collection, CINAHL. We also searched grey literature including OpenGrey, ProQuest Dissertations and conference proceedings of radiology and radiography associations in Nigeria. We checked reference lists of included studies. Search terms combined (Radiographer\* OR Medical imaging technologist\*) AND (Image interpretation OR Reporting OR PIE OR PCE OR Red Dot) AND (Training OR Education OR Competency) AND (Diagnostic OR Accuracy).

Selection process: We exported records to EndNote 20 and Rayyan. We removed 312 duplicates automatically and manually. Two reviewers independently screened titles, abstracts and full texts. Disagreements were resolved by discussion and a third reviewer made final decision where needed.

Data extraction: We used a standard Excel form and piloted it on 3 studies. We collected study characteristics, training details using TIDieR checklist (Hoffmann et al., 2014) with 12 items about type, reason, materials, content, duration, mode, location, trainer, image bank and reference standard (number of images, anatomy, normal/abnormal mix, validation, viewing conditions, single or consensus reference), and outcomes (baseline, post-training, change,

sensitivity, specificity, accuracy, predictive values, area under curve (AUC), Jackknife alternative free-response receiver operating characteristic (JAFROC), localization, description, confidence, false negatives/positives, p-values, confidence intervals, follow-up).

Risk of bias: We checked quality of all 23 studies. We used RoB2 (Sterne et al., 2019) for the RCT, ROBINS-I (Sterne et al., 2016) for 13 pre-post studies and QUADAS-2 (Whiting et al., 2011) for 9 audits. Two reviewers checked each study independently after piloting. Disagreements were resolved by consensus or third reviewer. Each study was rated as Low, Moderate/Some concerns, Serious or High risk. We showed results as traffic light plot (Green=Low, Yellow=Moderate/Some concerns, Orange=Serious, Red=High) and summary bar chart. We also checked if removing Serious/High risk studies changed the overall direction.

Synthesis: The test banks were very different (25 to 16,483 images), prevalence (50% in tests vs real clinical mix), reference standards (single radiologist vs consensus of 3 vs expert panel) and training length (9 hours to 1 year) were different, and so meta-analysis was not possible. We used SWiM (Campbell et al., 2020). We grouped findings into five themes: characteristics, baseline, effect, comparison across countries and implications for Nigeria. We counted direction of effect and reported p-values and confidence intervals (Boon and Thomson, 2021). Subgroups were country, anatomy and training type.

Reporting bias and certainty: We compared methods and results and checked for protocols (Neep et al. protocol published). Funnel plot was not possible, so we assessed publication bias qualitatively (Sterne et al., 2011). We used GRADE (Guyatt et al., 2008; Murad et al., 2017) for accuracy, sensitivity, specificity and description skills, and we upgraded for large effect and dose-response relationship.

## Results

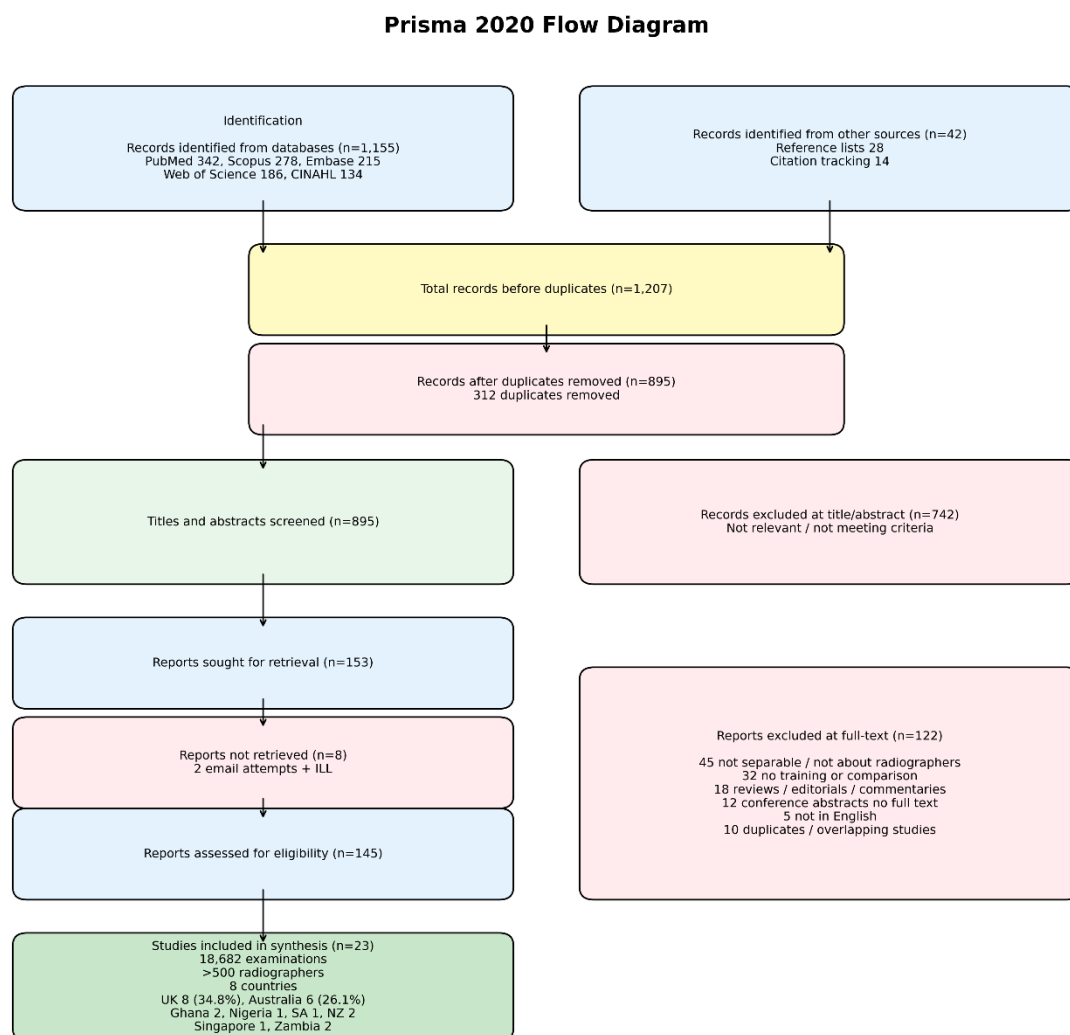


**Figure 1: PRISMA 2020 Flow Diagram. The study identified 1,207 records, 23 studies containing 18,682 examinations.**

Study selection: Database searching gave 1,155 records and other sources gave 42, total 1,207 before duplicates. After removing 312 duplicates, we screened 895 titles and abstracts and excluded 742. We tried to get 153 full texts and were able to assessed 145 for eligibility and excluded 122 (45 were not about radiographers or data could not be separated, 32 contains no training or comparison, 18 were reviews/editorials, 12 conference abstracts with no full text, 5 were not in English, and 10 duplicate or overlapping were found). Finally 23

studies were included.

Study characteristics: Eight studies were from UK (34.8%), six from Australia (26.1%), two from Ghana, one from Nigeria, one from South Africa, two from New Zealand, one from Singapore and two from Zambia. Thirteen studies (56.5%) were published since 2015, range 1994-2026. Design: one multi-centre stratified single-blind RCT (Neep et al., 2019 with 48 radiographers), 12 pre-post or quasi-experimental, 5 prospective/retrospective audits

and 5 diagnostic accuracy studies. Median sample size was 13 radiographers (range 7-121) but there were large audits with 2,402 cases (Alexander-Bates et al., 2021), 6,290 examinations (Brown et al., 2019), 11,090 audits (Takapautolo et al., 2024) and 16,483 images (Del Gante et al., 2021). In total we had 18,682 examinations and more than 500 radiographers. Settings: 15 emergency departments, 5 tertiary

hospitals, 2 rural district hospitals without radiologists (Munsanje, 2013). Median experience in African studies was 4 years and mean was 4.1 years (SD 5.2) in UK studies. Qualifications ranged from diploma (78% in Zambia) to Masters Level reporting radiographers. The study and the training characteristics were shown in table 1 and 2 respectively.

**Table 1: Study characteristics of 23 included studies.**

| Author/Year                     | Country      | Design                                   | Sample (Radiographers/Exams)                | Setting                               | Experience/Qualification                            |
|---------------------------------|--------------|------------------------------------------|---------------------------------------------|---------------------------------------|-----------------------------------------------------|
| Neep et al., 2019               | Australia    | Multi-centre stratified single-blind RCT | 48 radiographers                            | 6 public hospitals QLD                | Qualified radiographers, mixed experience           |
| Hazell et al., 2015             | South Africa | Pre-post interventional                  | 9 radiographers / 60 test cases             | Tertiary hospital                     | General radiographers                               |
| Ofori-Manteaw & Dzidzornu, 2019 | Ghana        | Action research pre-post                 | 8 radiographers / 30 skeletal               | District hospital                     | General radiographers                               |
| Ofori-Manteaw et al., 2026      | Ghana        | Quasi-experimental pre-post              | 36 radiographers / 60-image validated tool  | 3 hospitals (Ho, Kulmasa, Cape Coast) | General radiographers, junior/senior mix            |
| Ekpo et al., 2015               | Nigeria      | Cross-sectional diagnostic accuracy      | 51 radiographers / 50 chest CXRs (23N/27Ad) | Tertiary teaching hospitals           | Median >5 years predictive p=0.005, private p=0.004 |
| Brown et al., 2019              | Australia    | 12-month prospective longitudinal audit  | 12 radiographers / 6,290 examinations       | Emergency Department                  | PIE-trained, 100/month random audit                 |
| Lewis et al., 2024a             | New Zealand  | Pilot pre-post intensive 9h              | 7 radiographers / 60 extremity exams        | District health board                 | 7/28 volunteers (25% interest bias)                 |
| Lewis et al.,                   | New          | 6-month                                  | 10 radiographers                            | District                              | PIE-trained                                         |

|                           |           |                                         |                                            |                                   |                                                       |
|---------------------------|-----------|-----------------------------------------|--------------------------------------------|-----------------------------------|-------------------------------------------------------|
| 2024b                     | Zealand   | prospective PIE                         |                                            |                                   |                                                       |
| McConnell et al., 2012    | Australia | Pre-post condensed MSK                  | 10 radiographers                           | Rural/urban                       | General radiographers                                 |
| Smith et al., 2009        | Australia | Pre-post continuing education           | 16 radiographers                           | Rural hospitals                   | Rural general radiographers                           |
| Hargreaves & Mackay, 2003 | UK        | Pre-post Red Dot 8-week                 | 7 radiographers                            | A&E                               | General, very small sample                            |
| Williams et al., 2007     | UK        | Pre-post Bradford Course                | 121 radiographers                          | Bradford Trust                    | General radiographers, largest UK course              |
| Del Gante et al., 2021    | Singapore | Pre-post blended 2-month                | 13 radiographers / 16,483 images           | Tertiary                          | General radiographers                                 |
| Piper et al., 2005        | UK        | Diagnostic accuracy OSE                 | 28 reporting radiographers / 6,796 cases   | University + NHS                  | Postgraduate certificate, reporting radiographers     |
| Piper et al., 2014        | UK        | Diagnostic accuracy chest               | 40 reporting radiographers / 4,000 reports | NHS Trusts                        | Masters-level reporting radiographers                 |
| Woznitza et al., 2018     | UK        | JAFROC observer study                   | 11 reporting radiographers                 | Tertiary                          | Reporting radiographers vs 10 consultant radiologists |
| Harcus et al., 2023       | UK        | Pre-post online tutorials               | 9 radiographers / mean 4.1±5.2y            | ED                                | General, online recorded tutorials                    |
| Loughran, 1994            | UK        | Retrospective audit                     | 3 radiographers                            | A&E                               | Reporting radiographers, n=3 small                    |
| Robinson et al., 1999     | UK        | Retrospective audit selective reporting | 2 radiographers / 11,000 cases             | A&E reporting rota                | Highly experienced, selective                         |
| Munsanje, 2013            | Zambia    | Feasibility thesis rural                | Rural radiographers / Western Province     | District hospitals no radiologist | Diploma, rural                                        |
| Alexander-Bates et al.,   | Australia | 2-year retrospective                    | Multiple radiographers / 2,402             | ED                                | PIE-trained, false-negative focus                     |

|                          |           |                                  |                                 |                                  |                                                |
|--------------------------|-----------|----------------------------------|---------------------------------|----------------------------------|------------------------------------------------|
| 2021                     |           | e audit FN                       | cases                           |                                  |                                                |
| Takapautolo et al., 2024 | Australia | 5-year retrospective audit FP    | Multiple / 11,090 audits        | ED                               | PIE-trained, false-positive focus, 2.8% unsure |
| Chilambe et al., 2024    | Zambia    | Pre-post interventional WhatsApp | 27 radiographers / trauma chest | Copperbelt 8 hospitals Level 1-3 | Median 4y, 78% diploma, HPCZ registered        |

**The study characteristics of the 23 studies by country, design, sample size, setting, and experience.**

**Table 2: TIDieR intervention characteristics**

| Study             | Type                   | Duration                                               | Mode                      | Content Focus                                                            | Trainer                                | Image Bank                            | Reference Standard       |
|-------------------|------------------------|--------------------------------------------------------|---------------------------|--------------------------------------------------------------------------|----------------------------------------|---------------------------------------|--------------------------|
| Neep 2019         | PIE                    | 13.5h intensive 2-day vs 90min x9 weekly non-intensive | Face-to-face workshops x9 | Hand/wrist/fore arm, foot/ankle, systematic search, missed abnormalities | Reporting radiographers + radiologists | 60 DICOM trauma extremities validated | Radiologist panel + OSE  |
| Hazell 2015       | PIE commenting         | 6 weeks tailored + logbook                             | Face-to-face + mentoring  | Pattern recognition + comment construction                               | Radiologist + reporting radiographer   | 60 test cases                         | Radiologist report       |
| Ofori-Mantey 2019 | PIE trauma             | Short course                                           | Face-to-face              | Appendicular trauma skeleton                                             | Radiologist                            | 30 skeletal radiographs               | Radiologist              |
| Ofori-Mantey 2026 | PIE trauma             | 2-day intensive                                        | Face-to-face              | Anatomy, pathology, search, description, diagnosis formulation           | Radiologist                            | 60-image validated tool 30N/30Ab      | Gold standard panel      |
| Ekpo 2015         | Baseline (no training) | N/A                                                    | N/A - baseline assessment | Chest COPD, pneumonia, TB, effusion, pneumothorax                        | N/A                                    | 50 chest CXRs 23N/27Ab                | Consensus 3 radiologists |
| Brown             | PIE                    | Mandatory                                              | Online +                  | General ED                                                               | Radiologist                            | 6,290                                 | Radiologist              |

|                      |                         |                                                                                                       |                                                                            |                                                                                                         |                                        |                                                                           |                                                                                             |
|----------------------|-------------------------|-------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|----------------------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|
| 2019                 |                         | y 17-<br>module +<br>12-month<br>audit                                                                | clinical                                                                   | trauma                                                                                                  | ts +<br>reporting<br>radiograph<br>ers | clinical<br>examinati<br>ons                                              | ist report<br>+<br>100/mon<br>th<br>random<br>audit                                         |
| Lewis<br>2024a       | PIE                     | 9h<br>intensive<br>2-day                                                                              | Face-to-<br>face                                                           | Extremity<br>trauma 3Ws<br>(What Where<br>How)                                                          | MHSc<br>expert<br>facilitators         | 60<br>extremity<br>exams,<br>90sec/cas<br>e                               | Senior<br>consulta<br>nt review<br>discorda<br>nce                                          |
| Del<br>Gante<br>2021 | PIE<br>appendic<br>ular | 2-month<br>experienti<br>al +<br>blended                                                              | Seminars<br>+ online +<br>audit                                            | Appendicular<br>skeleton                                                                                | Radiologis<br>ts                       | 16,483<br>clinical<br>images                                              | Radiolog<br>ist report                                                                      |
| Piper<br>2005        | Skeletal<br>reporting   | 1-year<br>postgradu<br>ate<br>certificate                                                             | Academic<br>+ clinical                                                     | Appendicular +<br>axial skeleton                                                                        | Consultant<br>radiologist<br>s         | 6,796<br>OSE cases                                                        | Consulta<br>nt<br>radiologi<br>sts OSE                                                      |
| Piper<br>2014        | Chest<br>reporting      | Postgradu<br>ate<br>program<br>me                                                                     | Academic<br>+ clinical                                                     | Chest<br>pathology                                                                                      | Consultant<br>radiologist<br>s         | 4,000<br>reports                                                          | Expert<br>thoracic<br>panel                                                                 |
| Woznit<br>za 2018    | Chest<br>reporting      | Postgradu<br>ate<br>reporting                                                                         | Academic<br>+ clinical                                                     | Chest X-ray                                                                                             | Reporting<br>radiograph<br>ers         | Observer<br>study                                                         | Expert<br>thoracic<br>panel +<br>JAFROC                                                     |
| Harcus<br>2023       | PCE chest<br>trauma     | Online<br>recorded<br>tutorials<br>self-<br>paced<br>mean 84<br>days                                  | Online                                                                     | Systematic<br>search strategy,<br>how to structure<br>PCE, practice<br>cases                            | Reporting<br>radiograph<br>ers         | Test bank                                                                 | Radiolog<br>ist                                                                             |
| Chilam<br>be<br>2024 | Trauma<br>chest PIE     | 30-min<br>lecture +<br>8 weeks<br>WhatsApp<br>p 1<br>CXR/wee<br>k Mon<br>10:00<br>report Fri<br>15:00 | Blended:<br>lecture +<br>asynchron<br>ous<br>WhatsApp<br>+ Zoom<br>debrief | ABCDEF<br>systematic<br>(Airway,<br>Bones,<br>Circulation,<br>Diaphragm,<br>Everything else,<br>Fields) | Consultant<br>radiologist<br>s         | Hardcopy<br>film-<br>screen +<br>archive<br>cases, 2h<br>pre/post<br>test | Radiolog<br>ist panel<br>(consulta<br>nt<br>radiologi<br>sts Ndola<br>Teaching<br>Hospital) |

**The TIDieR intervention characteristic descriptions of training type, duration, mode and content.**

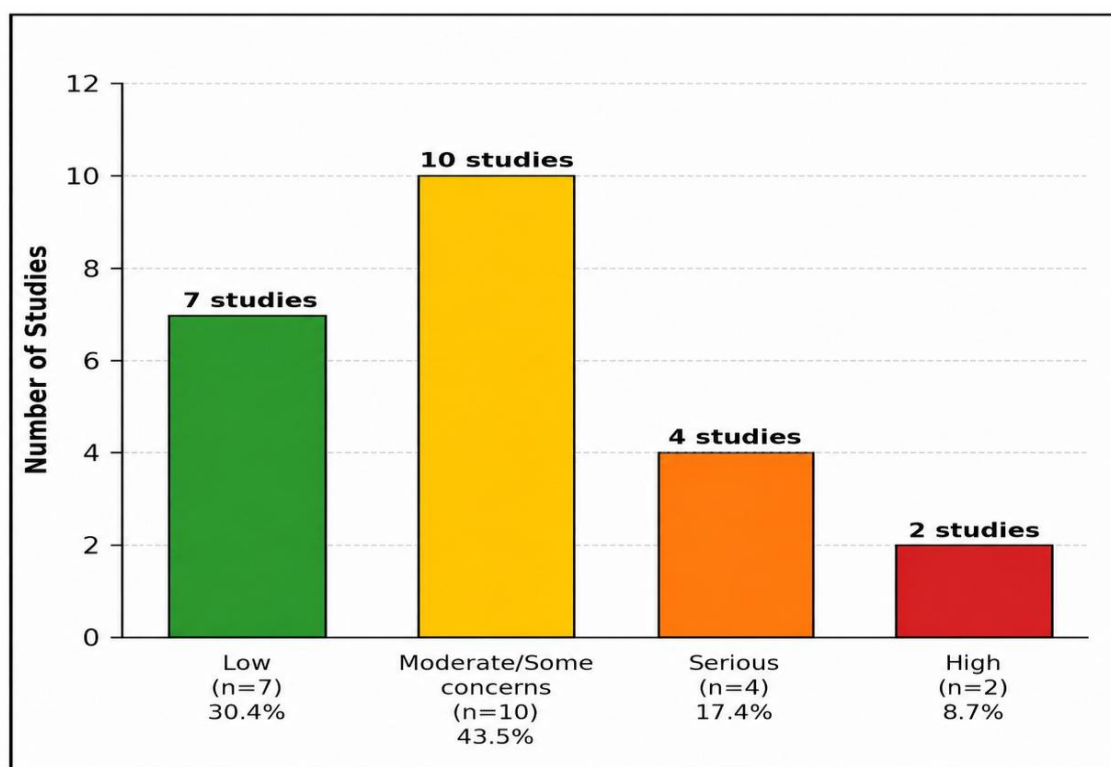


**Figure 2: Risk of bias summary. The counts indicated that 7 of the studies recorded low risk of bias, 10 moderate/some concerns, 4 serious and 2 showed high risk of bias.**

Risk of bias: For the RCT (Neep et al., 2019) we used RoB2 and found some concerns because 6 participants (12.5%) dropped out and all were from non-intensive arm (25% vs 0%). For 13 pre-post studies we used ROBINS-I. Nine had moderate risk (Hazell et al., 2015 100; Piper et al., 2005; Piper et al., 2014; Woznitza et al., 2018; Alexander-Bates et al., 2021; Takapautolo et al., 2024) and 2 had high risk (Robinson et al., 1999; Loughran, 1994). This is visualized in figure 2 and table 4. When we removed Serious/High studies, direction of findings did not change.

Baseline competency before training: Seven studies reported baseline. In Ghana accuracy was 68.8% (Ofori-Manteaw and Dzidzornu, 2019) and 71.7% (Ofori-Manteaw et al., 2026). In South Africa accuracy was 71.04%, sensitivity 83.73%, specificity 59.62% and incorrect comments 24.1% (Hazell et al., 2015). The Nigerian study (Ekpo et al., 2015) had 51 radiographers and 50 chest X-rays (23 normal, 27 abnormal including COPD, pneumonia, TB,

effusion, pneumothorax) compared with consensus of 3 radiologists. It found sensitivity 76.9% (95% CI 0.658-0.864), specificity 79.8% and false positive rate 20.2%. Experience more than 5 years was better ( $p=0.005$ ) and private practice also helped ( $p=0.004$ ). In UK and Australia baseline was also moderate, for example Red Dot sensitivity 0.71 (95% CI 0.62-0.79) with specificity 0.95-0.96 (Hargreaves and Mackay, 2003) and musculoskeletal description 89.9% (McConnell et al., 2012). Overall baseline without training was below 90% threshold needed for safe independent practice. Table 5 and figure 4 (supplementary) show baseline and post training results.

Effect of training on competency: Nineteen studies looked at training effect. All showed improvement in same direction and 15 were statistically significant. Accuracy in Ghana improved from 68.8% to 83.3% (+14.5%,  $p=0.017$ ) (Ofori-Manteaw and Dzidzornu, 2019) and from 71.7% to 79.5% (+7.8%,  $p=0.002$ ) (Ofori-Manteaw et al., 2026). In South Africa it

improved from 71.04% to 78% (+6.96%,  $z=2.66$ ,  $p=0.008$ ) (Hazell et al., 2015). In Australia condensed MSK training improved from 89.9% to 95% (McConnell et al., 2012). In New Zealand pilot it improved from 77.55% to 80.87% (up to 100%) and from 91.9% to 96.5% over 6 months with feedback.

**Red Dot and commenting:** UK Red Dot sensitivity improved from 0.71 to 0.81 with specificity 0.95-0.96 maintained (Hargreaves and Mackay, 2003). In Bradford Course with 121 radiographers Red Dot improved +9.8% (95% CI 2.4-17.1,  $p=0.012$ ) and commenting +12.7% (95% CI 3.8-21.5,  $p=0.007$ ) (Williams et al., 2007). Commenting was 13.7% less accurate than Red Dot before training and 10.8% less after, showing describing is harder than just flagging.

**Sensitivity, specificity and other measures:** In Ghana, sensitivity increased from 69.2% to 83.3% ( $p=0.042$ ) and specificity 68.3% to 83.3% ( $p=0.011$ ) (Ofori-Manteaw & Dzidzornu, 2019). The second study in Ghana (Ofori-Manteaw et al., 2026) recorded sensitivity 84.1% to 90.4% ( $p=0.012$ ), PPV 76.4% to 89.6% ( $p<0.001$ ) and AUC 0.821 to 0.878 ( $p=0.002$ ). South Africa sensitivity 83.73% to 87.28% and specificity 59.62% to 70.34%.

**Localisation and description skills had largest gains:** In Ghana, (Ofori-Manteaw et al., 2026) localization rose from 59.2% to 68.4% ( $p=0.003$ ), description 55.3% to 77.1% ( $p<0.001$ , +21.8%) and diagnosis 43.1% to 49.3% ( $p=0.022$ ). South Africa incorrect comments reduced 24.1% to 17.78% ( $p=0.05$ ). Singapore blended 2-month training with 16,483 images had sensitivity 87.4-98.9% mean 94.3%, specificity 96.4-99.9% mean 98.2% and accuracy 93.1-99.5% mean 96.9% (Del Gante et al., 2021). The Australian RCT (Neep et al., 2019) with 48 radiographers showed intensive 2-day 13.5-hour training was significantly better

than weekly 90-minute x9 training at 1-week ( $p=0.002$ ) and 12-week ( $p<0.001$ ) and confidence also improved.

**Sustainability:** Brown et al 12-month audit of 6,290 examinations showed accuracy 92.0%, sensitivity 71.1% and specificity 98.4% was sustained with 100 cases per month random audit (Brown et al. 2019). Lewis et al also showed 91.9% to 96.5% over 6 months with regular feedback (Lewis et al. 2024a).

**Error patterns:** Alexander-Bates et al 2-year audit of 2,402 cases with 88.7% accuracy found distal extremities especially phalanges of hand and feet accounted for 41.1% of missed pathology, and multi-injury accounted for 17 out of 60 false-negatives (28.3%) (Alexander-Bates et al. 2021). Takapautolo et al 5-year audit of 11,090 audits found accuracy 87.7%, sensitivity 79.1%, specificity 95.8% with 2.8% unsure cases, showing tendency to overcall (Takapautolo et al. 2024). This is visualized in figure 4.

**Comparison across countries and GRADE:** High-resource models UK (8 studies), Australia (6), Singapore (1) and New Zealand (2) achieved 90-96.9% for PIE and over 95% for full reporting through 1-year postgraduate certificate. However these need high cost. African studies achieved 79.5-83.3% after short 2-day to 6-week courses, showing low-cost models can work. GRADE certainty was Low for diagnostic accuracy (15 studies), sensitivity (18) and specificity (18) due to serious risk of bias, inconsistency (68.8-96.9% range), indirectness, imprecision (small African  $n=8-36$ ) and publication bias, but upgraded for large effect +14.5% and dose-response intensive better  $p<0.001$ ; Moderate for localisation/description (5 studies) due to large effect +21.8% and consistency. Table 3 and table 6 showed results on cross country comparison and GRADE certainty of the studies.

**Table 3: Cross-country comparison and scalability gradient**

| Region                                    | Accuracy                                                                                                                                                       | Training Model                                                                                                         | Cost/Resource                                                                 | Transferability to Nigeria                                                              | Policy Implication                                                                                                                                         |
|-------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| UK (n=8)                                  | PIE 90-96.9%, Reporting >95% sens/spec/acc 6,796 cases, chest 95.4% sens, JAFROC non-inferior FOM 0.828 vs 0.788 p<0.0001                                      | 1-year postgraduate certificate PgCert/PgDip Masters-level, mandatory 17-module, 100/month random audit, HCPC          | High cost, high resource, academic+ clinical, consultant radiologist trainers | Low direct transferability to district hospitals, high for tertiary centres             | Long-term: Establish accredited postgraduate reporting radiographer programme via RRBN, College of Radiographers model, Masters-level                      |
| Australia (n=6)                           | PIE 88.7-92.0% acc 71.1% sens 98.4% spec sustained, FN phalanges 41.1%, FP 87.7% sens 79.1% spec 95.8% 2.8% unsure                                             | Intensive 2-day 13.5h > non-intensive weekly p=0.002 1-week p<0.001 12-week, condensed MSK, 12-month prospective audit | Moderate cost, 17-module mandatory, online + clinical                         | Moderate-high: Intensive 2-day model feasible for CPD, audit model 100/month applicable | Adopt 100/month random audit per Brown, focus curriculum on distal extremities phalanges and multi-injury (Alexander-Bates) and unsure cases (Takapautolo) |
| Singapore (n=1) / NZ (n=2)                | 94.3% sens 98.2% spec 96.9% acc 16,483 images, NZ 90-96.9% comparable UK/Australia, pilot 77.55%→80.87% up to 100% and 91.9%→96.5% over 6 months with feedback | 2-month blended seminars+online+audit (Singapore), 9h intensive 2-day 3Ws What Where How + 6-month feedback (NZ)       | Moderate, blended feasible                                                    | High: Blended and intensive with feedback shows transferability to new settings         | Use blended + feedback model: 2-day intensive + 6-month feedback to sustain 91.9%→96.5%                                                                    |
| Africa Ghana 2, SA 1, Zambia 2, Nigeria 1 | Baseline 68.8-71.7% Ghana, 71.04% SA, Nigeria location sens 88.9% sens 76.9% spec 79.8%; Post 79.5-83.3% Ghana,                                                | Short 2-day intensive, 6-week tailored + logbook, 30-min lecture ABCDEF + 8 weeks WhatsApp 1                           | Low cost, low resource, smartphone penetration >80%, hardcopy film-screen     | Very high: Chilambe Zambia model most transferable - 8 hospitals Level 1 District to    | Immediate: Phase 1 Red Dot triage (Munsanje rural), Phase 2 PIE trauma appendicular distal extremities (Ghana 2-day +14.5% p=0.017), Phase 3 Chest TB      |

|                                                                                                                                                     |                                                                               |                           |                                                                                                                                               |                                                                                                                                                                                                                                                                                     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|---------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 78% SA, Zambia existing knowledge 92-96% sensitivity literature but improvement post-training; Nigeria years >5y p=0.005 private p=0.004 predictive | CXR/week Mon 10:00 report Fri 15:00 data allowances + Zoom debrief (Chilambe) | feasible, data allowances | Level 3 Teaching, median 4y 78% diploma HPCZ registered similar Nigeria <5,000 radiographers diploma-heavy, district hospitals no radiologist | reporting tertiary; Implement WhatsApp mentoring with data allowances, ABCDEF systematic, 60-image validated tool 50% prevalence, 90sec/case, 3Ws; Policy: RRBN define scope Red Dot→PIE→full reporting, indemnity, referral pathways, stakeholder engagement RRBN, ARN, MDCN, FMOH |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|---------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

The table shows the UK/Australia models (high-resource) versus African models (low-resource) and their scalability.

## Discussion

This review of 23 studies with 18,682 examinations and more than 500 radiographers shows that baseline skill without training is moderate (68.8-76.9%) below 90% safety threshold, but short low-cost training improves it significantly to 79.5-83.3% in Africa and up to 92.0-96.9% in Australia/Singapore/New Zealand, with full reporting over 95% in UK after postgraduate programme.

Baseline findings were similar in all countries. This means moderate performance without formal training is a global issue, not only African. The Nigerian chest study showed reasonable starting point, especially location sensitivity 88.9%, and supports starting with experienced radiographers. This matches finding that more than 5 years experience was better.

Training effect was consistent in direction in all 19 studies and 15 were statistically significant. The biggest gain was for description skills +21.8% (Ofori-Manteaw et al., 2026). This means curriculum should focus on What-Where-How reporting. The RCT (Neep et al., 2019) gives best evidence that concentrated intensive training works better than weekly short sessions. Sustainability data show one-off training is not

enough and error analysis gives clear guidance for curriculum (Alexander-Bates et al., 2021). Phalanges of hand and feet caused 41.1% of missed fractures and multi-injury caused 28.3% of misses (Alexander-Bates et al., 2021; Takapautolo et al., 2024; Brown et al. 2019).

Comparison across countries shows a scalability gradient from expensive 1-year academic programmes to cheap WhatsApp mentoring shown in table 3. UK/Australia models achieve highest accuracy but need resources not very available in Nigerian district hospitals. The Zambian model by Chilambe et al. (2024) with 27 radiographers median 4 years 78% diploma in 8 hospitals Level 1 District to Level 3 Teaching hospitals using 30-minute lecture on ABCDEF plus 8 weeks asynchronous WhatsApp with 1 chest X-ray per week (Monday 10:00 training, Friday 15:00 discussion via text/voice note) with data allowances is most transferable. Singapore blended 2-month model with 16,483 images achieving mean 96.9% accuracy and Australian online recorded tutorials improving specificity from 79.0% to 89.9% with mean 84 days self-paced (Harcus et al., 2023) are also scalable across six geopolitical zones in Nigeria.

Strengths of this review: First, it is first to focus

on Nigeria and sub-Saharan Africa training gap and show lack of data compared to other countries. Second, it proposes a phased implementation model for Nigeria starting with

most transferable low-cost Zambian model. Third, it uses many quality tools – PRISMA 2020, SWiM, TIDieR, RoB2, ROBINS-I, QUADAS-2 and GRADE.

## Nigeria Implementation Model

**Figure 6: Phased Implementation Model for Workforce Extension in Nigeria Based on Systematic Review Findings (n=23 studies, 18,682 examinations)**

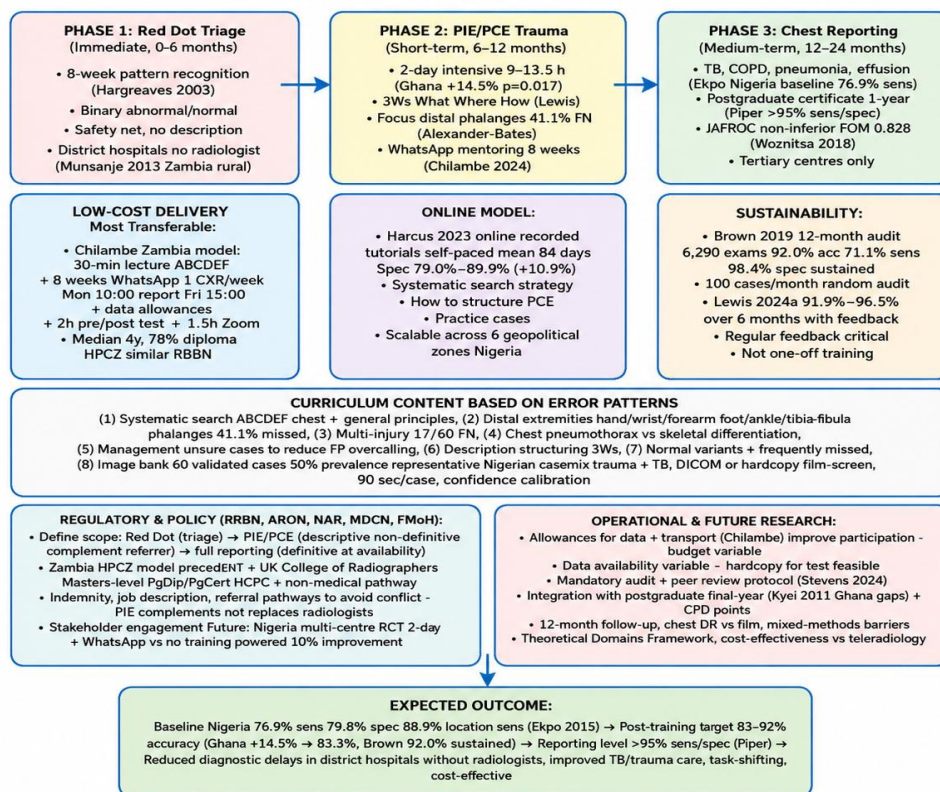


**Figure 3: Phased implementation model for workforce extension in Nigeria based on the systematic review of the 23 studies.**

We propose phased approach. Phase 1 Red Dot triage immediate 0-6 months 8-week pattern recognition for district hospitals without radiologists (similar to Munsanje, 2013 Zambia rural model). Phase 2 PIE/PCE trauma short-term 6-12 months 2-day intensive 9-13.5-hour training (Ghana model +14.5% p=0.017) with 3Ws What Where hoW (Lewis et al.), focusing on distal phalanges where 41.1% errors occur more. Delivery should use WhatsApp mentoring like Chilambe et al. (2024) model with data allowances, plus online recorded tutorials (Harcus et al., 2023) which improved specificity

79.0% to 89.9%. Phase 3 chest reporting medium-term 12-24 months tertiary centres for TB, COPD, pneumonia, effusion, pneumothorax etcetera building on Ekpo Nigerian baseline, needs 1-year postgraduate certificate (Piper et al., >95%) and JAFROC non-inferiority evidence (Woznitza et al., 2018). This can serve as good starting model with gradually transition to whole/general skeletal formal image reporting ranging from head to toe using the UK model as a specialist certification for reporting radiographers. This will inevitably ensure timely and accurate report where the radiologist are not

available and complement radiologists role of reporting for timely reporting of A/E cases and for routine preliminary reporting in projection (conventional) imaging practices.

To keep skills, mandatory audit of 100 cases per month random audit against radiologist report (Brown et al., 2019) and regular feedback is required. Allowances for data and transport improve participation and should be budgeted. This can come as a specialist allowance for experience radiographers who met the criteria for the role. Digital radiography (DR) availability is variable so hardcopy film-screen for test is feasible. Regulatory bodies RRBN, MDCN and Federal Ministry of Health need to define scope clearly backed with legislations – Red Dot (triage), PIE/PCE (descriptive non-definitive), Full reporting (definitive accredited) with indemnity and referral pathways to avoid conflict, drawing on Zambia HPCZ and UK College of Radiographers Masters-level models. Integration into undergraduate final-year (Kyei et al., 2011) and CPD points is important.

The study has huge gap of publication among African countries. More researches are desirable. Future research in Nigeria should include multi-centre RCT of 2-day intensive plus WhatsApp mentoring versus no training powered for 10% improvement with 12-month follow-up, chest DR versus film comparison, mixed-methods barriers using Theoretical Domains Framework, and cost-effectiveness versus tele-radiology.

**Limitations:** First, there was serious heterogeneity in test banks, reference standards, viewing conditions and training duration (9 hours to 1 year), so meta-analysis was not possible and we used SWiM. Second, risk of bias was moderate to serious in 15 out of 23 studies. However sensitivity analysis excluding Serious/High risk studies did not change direction, and GRADE was Low for accuracy/sensitivity/specificity and Moderate for description due to large effect +21.8% and dose-response.

## Conclusion

Baseline competency without training is moderate and below safe threshold, but short

low-cost training significantly improves it globally and more significantly with postgraduate program. The Zambian WhatsApp model is most transferable to Nigerian district hospitals without radiologists. Gradual Phased implementation incorporating Red Dot, Preliminary commenting to definitive reporting role in projection imaging, with mandatory regular audit and feedback, offers feasible, scalable, cost-effective workforce extension to reduce diagnostic delays and errors in Nigeria.

## Implications for practice

Short intensive training combined with low-cost WhatsApp mentoring can improve radiographer image interpretation from moderate baseline to good performance, especially in district hospitals without radiologists. Starting with the Zambian model and gradual to whole skeletal reporting in projection imaging with regular audit and feedback, offers a feasible pathway. Clear regulation, job descriptions, indemnity, budget for allowances and supervision from Radiographers Registration Board of Nigeria and other stakeholders are required.

## References

- Alexander-Bates I, Neep MJ, Starkey D, et al. An analysis of radiographer preliminary image evaluation – A focus on common false negatives. *J Med Radiat Sci*. 2021;68(2):155-162.
- Berman L, de Lacey G, Twiston-Davies E, et al. Reducing errors in the accident department: a simple method using radiographers. *Br Med J (Clin Res Ed)*. 1985;290:421-422.
- Brady AP. Error and discrepancy in radiology: inevitable or avoidable? *Insights Imaging*. 2017;8:171-182.
- Bramer WM, Rethlefsen ML, Kleijnen J, Franco OH. Optimal database combinations for literature searches in systematic reviews. *Syst Rev*. 2017;6:245.
- Brealey S, Scally AJ, Thomas NB. Methodological standards in radiographer plain radiograph reporting in clinical practice: a meta-analysis. *Clin Radiol*. 2005;60:232-241.



Brown C, Neep MJ, Pozzias E, Starkey D. Reducing risk in the emergency department: a 12-month prospective longitudinal study of radiographer preliminary image evaluations. *J Med Radiat Sci.* 2019;66:26-32.

Brunswick JE, Ilkhanipour K, Seaberg DC, Brown LH. Accuracy of radiographic readings in the emergency department. *Am J Emerg Med.* 2004;22:320-323.

Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ.* 2020;368:l6890.

Croskerry P. Clinical cognition and diagnostic error. *Adv Health Sci Educ.* 2009;14:27-35.

Del Gante E, et al. Accuracy of radiographer comment following a two-month experiential and blended learning in appendicular skeleton X-ray interpretation: The Singapore experience. *Radiography.* 2021;27:433-439.

Ekpo EU, Egbe NO, Akpan BE, et al. Radiographers' performance in chest X-ray interpretation: the Nigerian experience. *Br J Radiol.* 2015;88:20150023.

Guly HR. Diagnostic errors in an accident and emergency department. *Emerg Med J.* 2001;18:263-269.

Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence. *BMJ.* 2008;336:924-926.

Hallas P, Ellingsen T. Errors in fracture diagnoses in the emergency department. *BMC Emerg Med.* 2006;6:4.

Hargreaves J, Mackay S. The accuracy of the Red Dot system. *Radiography.* 2003;9:101-107.

Hardy M, Snaith B, Smith T. Radiographer reporting of trauma images: United Kingdom experience. *Radiographer.* 2008;55:16-21.

Hardy M, Culpan G. The Bradford Red Dot Course. *Radiography.* 2007;13:124-129.

Hazell L, et al. Tailored pattern recognition and comment construction with logbook mentoring: South Africa. *Radiography.* 2015;21:340-345.

Higgins JPT, Thomas J, Chandler J, et al. *Cochrane Handbook for Systematic Reviews of Interventions.* Cochrane; 2022.

Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: TIDieR checklist. *BMJ.* 2014;348:g1687.

Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. *Cochrane Database Syst Rev.* 2011.

Landis JR, Koch GG. The measurement of observer agreement. *Biometrics.* 1977;33:159-174.

Loughran CF. Reporting of fracture radiographs by radiographers. *Br J Radiol.* 1994;67:945-950.

McConnell J, Devaney C, Gordon M, et al. The impact of a condensed education programme on accuracy of MSK trauma reporting by radiographers. *Radiography.* 2012;18:125-129.

McInnes MDF, Moher D, Thombs BD, et al. PRISMA-DTA Statement. *JAMA.* 2018;319:388-396.

Methley AM, et al. PICO, PICOS and SPIDER. *BMC Med Res Methodol.* 2014;14:121.

Munn Z, et al. Methodological guidance for systematic reviews of observational studies. *JBIM Evidence Synthesis.* 2020.

Murad MH, et al. Rating certainty in evidence in absence of single estimate. *J Clin Epidemiol.* 2017;87:10-18.

Murphy A, Ekpo E, Steffens T, Neep MJ. Radiographic image interpretation by Australian radiographers: systematic review. *J Med Radiat Sci.* 2019;66:269-283.

Munsanje F. Need for image reporting by radiographers in Zambia – Western Province rural hospitals. Thesis. 2013.

Neep MJ, Steffens T, Eastgate P, McPhail SM. Evaluating effectiveness of intensive versus non-intensive image interpretation education: RCT. *J Med Radiat Sci.* 2019;66:20-27.

Neep MJ, et al. Use of radiographer abnormality detection systems by Queensland public

hospitals. *J Med Radiat Sci.* 2014.

Ofori-Manteaw BB, et al. Assessing Ghanaian radiographers' competency in PIE: pre- and post-test study. *Radiography.* 2026;32:103224.

Okeji MC, et al. Factors influencing choice of radiology as specialty among medical students Nigeria. *BMC Med Educ.* 2023.

Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan. *Syst Rev.* 2016;5:210.

Page MJ, McKenzie JE, Bossuyt PM, et al. PRISMA 2020 statement. *BMJ.* 2021;372:n71.

Piper K, Paterson A. Initial image interpretation of appendicular skeletal radiographs. 2005.

Piper K, et al. Chest reporting by radiographers: accredited postgraduate programme. *Radiography.* 2014;20:94-99.

Pinto A, Brunese L. Spectrum of diagnostic errors in radiology. *World J Radiol.* 2010;2:377-383.

Radiographers Registration Board of Nigeria. Nigeria has less than 5,000 radiographers for over 200m population. *Trust Radio.* 2022.

Robinson PJA, et al. Radiographer reporting in A&E: 11,000 cases. *Clin Radiol.* 1999.

Royal College of Radiologists. Clinical Radiology Workforce Census 2023: 30% shortfall projected 40% by 2028. *RCR;* 2023.

Royal College of Radiologists. Clinical Radiology Workforce Census 2024: 29%

shortfall 39% by 2029 deficit 1,953. *RCR;* 2024.

Snaith BA, Hardy M, Lewis EF. Radiographer reporting in the UK: A longitudinal analysis. *Radiography.* 2015;21:119-123.

Smith T, et al. Influence of continuing education program on image interpretation accuracy of rural radiographers. *Rural Remote Health.* 2009.

Sterne JAC, Savović J, Page MJ, et al. RoB 2: revised tool for assessing risk of bias in randomised trials. *BMJ.* 2019;366:14898.

Sterne JAC, Hernán MA, Reeves BC, et al. ROBINS-I: tool for assessing risk of bias in non-randomised studies. *BMJ.* 2016;355:i4919.

Sterne JAC, et al. Recommendations for examining funnel plot asymmetry. *BMJ.* 2011;343:d4002.

Swinburne K. Pattern recognition among technically skilled radiographers. *Lancet.* 1971;2:131.

Takapautolo J, Neep M, Starkey D. Analysing false-positive errors when Australian radiographers use PIE. *J Med Radiat Sci.* 2024;71:540-546.

Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: revised tool for quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155:529-536.

Williams I, et al. Pilot pre-post Bradford Red Dot Course. *Radiography.* 2007;13:124-129.

Woznitza N, et al. Chest X-ray reporting by radiographers: JAFROC observer study. *Radiography.* 2018;24:146-152.

Berman, L., de Lacey, G., Twomey, E., Twomey, B., Welch, T. and Eban, R., 1985. Reducing errors in the accident department: a simple method using radiographers.

Brady, A.P., 2017. Error and discrepancy in radiology: inevitable or avoidable? *Insights into Imaging*, 8(1), pp.171–182.

Brealey, S. and Scally, A., 2008. Methodological approaches to evaluating Red Dot systems and radiographer reporting. *Radiography*, 14(1), pp.e11–e15.

Brealey, S., Scally, A., Hahn, S., Thomas, N., Godfrey, C. and Coomarasamy, A., 2005. Accuracy of radiographer plain radiograph reporting: a systematic review.

Brunswick, J.E., Ilkhanipour, K., Seaberg, D.C. and McGill, L., 2004. Radiographic interpretation in the emergency department.

Chilambe, F., et al., 2024. Radiographer image interpretation training using a WhatsApp

mentoring programme.

Hallas, P. and Ellingsen, T., 2006. Errors in fracture diagnosis in emergency departments.

Hardy, M. and Snaith, B., 2009. Role extension and image interpretation: an evolving radiographer role. *Radiography*, 15(4), pp.273–282.

Hardy, M., Culpan, G. and Snaith, B., 2013. Radiographer image interpretation: a review of the evidence. *Radiography*, 19(2), pp.e90–e98.

Hargreaves, J. and Mackay, S., 2003. The accuracy of the Red Dot system after image interpretation training.

Hazell, L., et al., 2015. Pattern-recognition training for radiographers.

Krupinski, E.A., 2010. Current perspectives in medical image perception. *Attention, Perception, & Psychophysics*, 72(5), pp.1205–1217.

Murphy, M., et al., 2019. Review of radiographer image interpretation and reporting.

Neep, M.J., Steffens, T., Eastgate, P. and McPhail, S.M., 2019. Evaluating the effectiveness of intensive versus non-intensive image interpretation education for radiographers: a randomised controlled trial. *Journal of Medical Radiation Sciences*, 66(1), pp.5–13.

Okeji, M.C., et al., 2023. Radiologist workforce in Nigeria.

Pinto, A. and Brunese, L., 2010. Spectrum of diagnostic errors in radiology.

Radiographers Registration Board of Nigeria (RRBN), 2022. Annual Report.

Snaith, B. and Hardy, M., 2008. Preliminary Clinical Evaluation by radiographers.

Society and College of Radiographers, 2022. Education and Career Framework for the Radiography Workforce. London: Society and

College of Radiographers.

Swinburne, K., 1971. Pattern recognition by radiographers.

The Royal College of Radiologists, 2023. Clinical Radiology UK Workforce Census Report 2023. London: The Royal College of Radiologists.

World Health Organization, 2023. Medical Imaging and Diagnostic Services. Geneva: World Health Organization.

## Declarations

**Funding:** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Conflict of interest:** Declarations of interest: none of the authors has conflict of interest to declare.

**Declaration of generative AI and AI-assisted technologies:** During preparation of this work the author(s) used Meta AI / ChatGPT to improve language clarity and formatting. After using this tool, the author(s) reviewed and edited the content as needed and take full responsibility for the published article.

**Data availability:** All data extracted are included in Tables 1-6. Search strategy and PRISMA checklist available as supplementary material.

**Word count:** Main manuscript approximately 3,920 words excluding abstract, references, tables and figures.

## Appendix

### Supplementary files

This document contains supplementary information to the main work.

**Table 4: Risk of Bias full traffic light for all 23 studies**

| # | Study                           | Country      | Tool     | Overall Judgement | Key Concerns / Domains                                                                                                                                                                                                               |
|---|---------------------------------|--------------|----------|-------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | Neep et al., 2019               | Australia    | RoB2     | Some concerns     | Missing data 6 dropouts 12.5% all non-intensive arm (25% vs 0%), interest/time bias; randomization low, deviations low, measurement blinded low, selection low                                                                       |
| 2 | Hazell et al., 2015             | South Africa | ROBINS-I | Moderate          | Confounding mixed experience no adjustment moderate, selection convenience moderate, classification low, deviations low, missing low, measurement low, reporting low                                                                 |
| 3 | Ofori-Manteaw & Dzidzornu, 2019 | Ghana        | ROBINS-I | Serious           | Serious confounding junior doctors co-intervention, serious selection small n=8 single site, moderate classification, low deviations, serious missing no follow-up, moderate measurement single radiologist, moderate reporting      |
| 4 | Ofori-Manteaw et al., 2026      | Ghana        | ROBINS-I | Moderate          | Moderate confounding experience, low selection 36 radiographers 3 hospitals, low classification validated tool, low deviations, low missing, low measurement gold standard panel, low reporting                                      |
| 5 | Ekpo et al., 2015               | Nigeria      | QUADAS-2 | Low               | Low patient selection (51 radiographers 50 chest), low index test (CXR interpretation), low reference standard consensus 3 radiologists, low flow/timing; foundational Nigerian study                                                |
| 6 | Brown et al., 2019              | Australia    | QUADAS-2 | Low               | Low patient selection 6,290 examinations, low index PIE, moderate reference single radiologist mitigated by 100/month random audit, low flow; 12-month prospective longitudinal                                                      |
| 7 | Lewis et al., 2024a             | New Zealand  | ROBINS-I | Serious           | Serious selection 7/28 volunteers 25% interest bias, serious confounding self-selection motivated, moderate classification, low deviations, moderate missing, low measurement, moderate reporting; pilot pre-post intensive 9h       |
| 8 | Lewis et al., 2024b             | New Zealand  | ROBINS-I | Moderate          | Moderate selection 10 radiographers district, moderate confounding experience, low classification, low deviations, low missing, low measurement senior consultant discordance review, low reporting; 6-month prospective 91.9%→96.5% |

|    |                           |           |          |          |                                                                                                                                                                                                                                                                  |
|----|---------------------------|-----------|----------|----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 9  | McConnell et al., 2012    | Australia | ROBINS-I | Moderate | Moderate confounding small n=10, moderate selection convenience, low classification condensed MSK programme, low deviations, low missing, low measurement radiologist reference, low reporting; 89.9%→95%                                                        |
| 10 | Smith et al., 2009        | Australia | ROBINS-I | Moderate | Moderate confounding rural mixed experience, moderate selection 16 rural radiographers, low classification continuing education, low deviations, low missing, low measurement, low reporting                                                                     |
| 11 | Hargreaves & Mackay, 2003 | UK        | ROBINS-I | Serious  | Serious confounding very small n=7 no control, serious selection historical control, serious classification 8-week pattern recognition, low deviations, serious missing, moderate measurement, moderate reporting; Red Dot sens 0.71→0.81                        |
| 12 | Williams et al., 2007     | UK        | ROBINS-I | Moderate | Low confounding large n=121 Bradford Course reduces bias, moderate selection single trust, low classification, low deviations, low missing, low measurement, low reporting; Red Dot +9.8% p=0.012 commenting +12.7% p=0.007                                      |
| 13 | Del Gante et al., 2021    | Singapore | ROBINS-I | Moderate | Moderate confounding 13 radiographers mixed experience, moderate selection single tertiary, low classification 2-month experiential+blended, low deviations, low missing, low measurement radiologist report, low reporting; 16,483 images sens 94.3% spec 98.2% |
| 14 | Piper et al., 2005        | UK        | QUADAS-2 | Low      | Low patient selection 6,796 OSE cases 28 radiographers, low index skeletal reporting, low reference consultant radiologists OSE, low flow; >95% sens/spec/acc postgraduate certificate                                                                           |
| 15 | Piper et al., 2014        | UK        | QUADAS-2 | Low      | Low patient selection 4,000 reports 40 radiographers, low index chest reporting, low reference expert thoracic panel, low flow; sens 95.4% CI 94.4-96.3% agreement 89%                                                                                           |
| 16 | Woznitza et al., 2018     | UK        | QUADAS-2 | Low      | Low patient selection 11 reporting radiographers observer study, low index chest X-ray reporting, low reference expert thoracic panel, low flow; JAFROC FOM 0.828 vs 0.788 radiologists p<0.0001 non-inferiority                                                 |
| 17 | Harcus et al., 2023       | UK        | ROBINS-I | Moderate | Moderate confounding small n=9 mean 4.1±5.2y, moderate selection ED, low classification online recorded tutorials mean                                                                                                                                           |

|    |                              |           |          |          |                                                                                                                                                                                                                                                                                                                                                   |
|----|------------------------------|-----------|----------|----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    |                              |           |          |          | 84 days self-paced, low deviations, low missing, low measurement radiologist, low reporting; spec 79.0%→89.9%                                                                                                                                                                                                                                     |
| 18 | Loughran, 1994               | UK        | QUADAS-2 | High     | High patient selection n=3 radiographers only 6 months fracture reporting, moderate index, high reference unclear single radiologist, high flow selective; early reporting radiographer study                                                                                                                                                     |
| 19 | Robinson et al., 1999        | UK        | QUADAS-2 | High     | High patient selection selective A&E reporting rota 2 radiographers only 11,000 cases highly experienced, moderate index, high reference selective, high flow selective rota; no significant clinical impact vs trainees                                                                                                                          |
| 20 | Munsanje, 2013               | Zambia    | ROBINS-I | Serious  | Serious confounding thesis rural no radiologist, serious selection Western Province rural convenience, serious classification unclear intervention, low deviations, serious missing, serious measurement unclear reference, serious reporting; feasibility need for reporting                                                                     |
| 21 | Alexander-Bates et al., 2021 | Australia | QUADAS-2 | Low      | Low patient selection 2,402 cases PIE-trained, low index PIE, low reference senior consultant discordance review, low flow; false-negative focus accuracy 88.7% phalanges 41.1% missed multi-injury 17/60                                                                                                                                         |
| 22 | Takapautolo et al., 2024     | Australia | QUADAS-2 | Low      | Low patient selection 11,090 audits 5-year largest FP analysis, low index PIE, low reference radiologist report, low flow; accuracy 87.7% sens 79.1% spec 95.8% 2.8% unsure overcalling                                                                                                                                                           |
| 23 | Chilambe et al., 2024        | Zambia    | ROBINS-I | Moderate | Moderate confounding 27 radiographers median 4y 78% diploma, moderate selection Copperbelt 8 hospitals Level 1-3, low classification 30-min ABCDEF + 8 weeks WhatsApp 1 CXR/week data allowances, low deviations, low missing, low measurement consultant radiologists Ndola Teaching Hospital panel, low reporting; most transferable to Nigeria |

**Table 4 contains full RoB2, ROBINS-1 and QUADAS-2 ratings for all studies.**

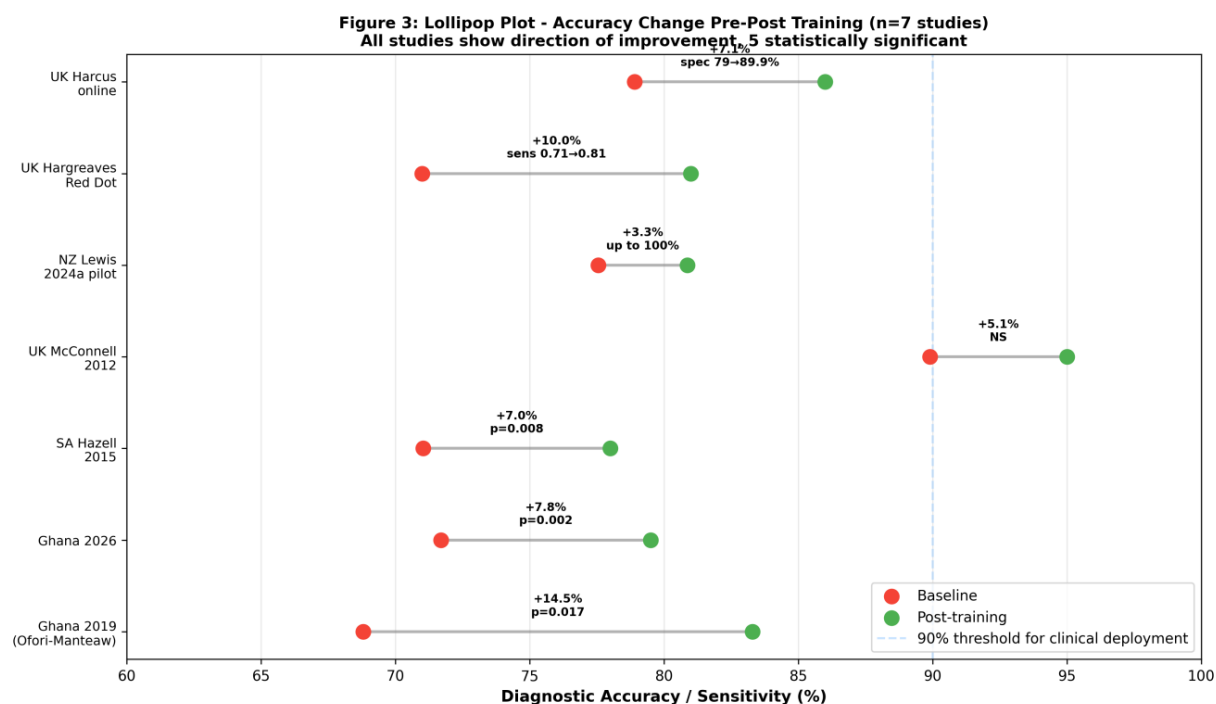


**Figure 4: Lollipop Plot of training effect.** This shows the direction and magnitude of improvement for 19 training studies.

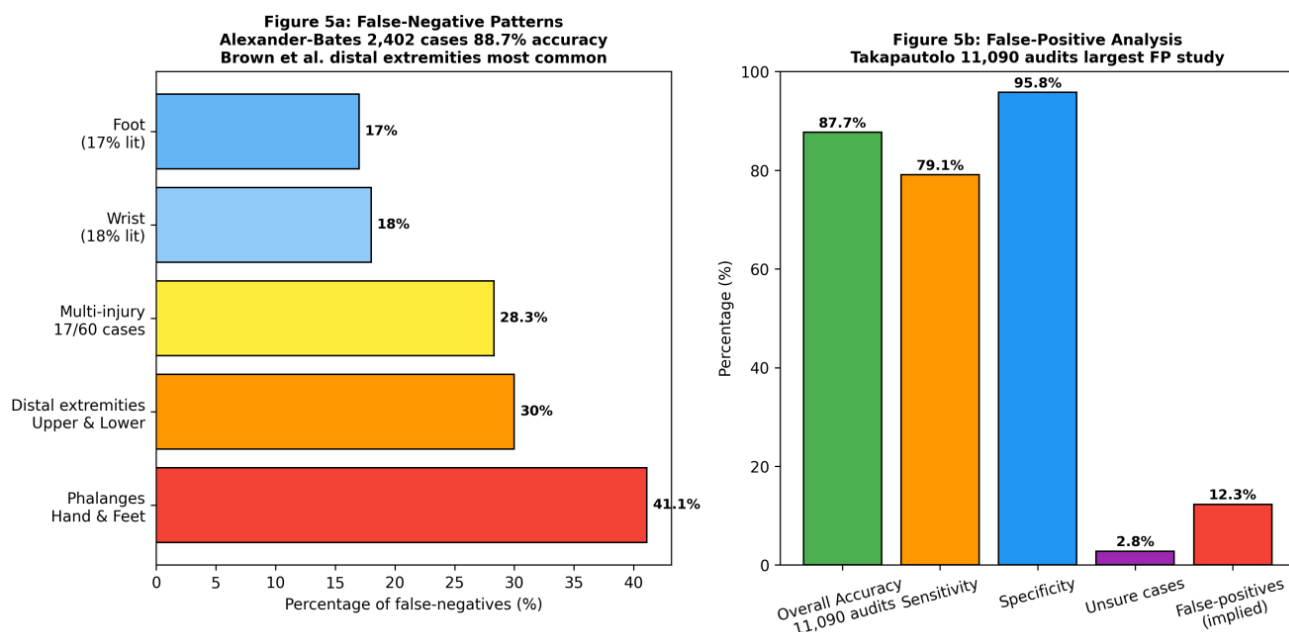


**Figure 5: Error patterns.** The magnitude of error in distal phalanges was 41.1% of false-negative, while cases of multi-injury was 28.3%.

**Table 5: Effect of training.**

| Theme      | Study                      | Baseline                                                                       | Post-Training                                                                | Change / p-value                                                                 |
|------------|----------------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------|----------------------------------------------------------------------------------|
| 1 Baseline | Ghana 2019 (Ofori-Manteaw) | 68.8% accuracy                                                                 | -                                                                            | Baseline moderate                                                                |
| 1 Baseline | Ghana 2026                 | 71.7% accuracy                                                                 | -                                                                            | Baseline moderate                                                                |
| 1 Baseline | South Africa Hazell 2015   | 71.04% acc, 83.73% sens, 59.62% spec                                           | -                                                                            | 24.1% incorrect comments                                                         |
| 1 Baseline | Nigeria Ekpo 2015          | Loc sens 88.9% CI 0.787-0.980, sens 76.9% CI 0.658-0.864, spec 79.8% FPR 20.2% | -                                                                            | >5y p=0.005, private p=0.004                                                     |
| 1 Baseline | UK Hargreaves 2003 Red Dot | Sens 0.71 CI 0.62-0.79, spec ~0.95                                             | -                                                                            | Below threshold 90%                                                              |
| 2 Effect   | Ghana 2019                 | 68.8% acc, 69.2% sens, 68.3% spec                                              | 83.3% acc, 83.3% sens, 83.3% spec                                            | +14.5% p=0.017, sens p=0.042, spec p=0.011                                       |
| 2 Effect   | Ghana 2026                 | 71.7% acc, 84.1% sens, 80% spec, PPV 76.4%, AUC 0.821, loc 59.2%, desc 55.3%   | 79.5% acc, 90.4% sens, 85% spec, PPV 89.6%, AUC 0.878, loc 68.4%, desc 77.1% | +7.8% p=0.002, sens p=0.012, PPV p<0.001, AUC p=0.002, loc p=0.003, desc p<0.001 |
| 2 Effect   | South Africa Hazell        | 71.04% acc, 83.73% sens, 59.62% spec                                           | 78% acc, 87.28% sens, 70.34% spec                                            | +6.96% z=2.66 p=0.008, incorrect comments 24.1%→17.78% p=0.05                    |
| 2 Effect   | Australia McConnell 2012   | 89.9% MSK                                                                      | 95%                                                                          | +5.1%                                                                            |
| 2 Effect   | NZ Lewis 2024a pilot       | 77.55% acc, 91.9% over time                                                    | 80.87% acc (+3.33% up to 100%), 96.5% over 6 months                          | Sustained with feedback                                                          |

|                   |                         |                                 |                                                                                               |                                                         |
|-------------------|-------------------------|---------------------------------|-----------------------------------------------------------------------------------------------|---------------------------------------------------------|
| 2 Effect          | UK Hargreaves Red Dot   | Sens 0.71                       | Sens 0.81                                                                                     | +0.10 sens, spec maintained 0.95-0.96                   |
| 2 Effect          | UK Williams Bradford    | Red baseline Dot                | Red Dot +9.8% CI 2.4-17.1 p=0.012, commenting +12.7% CI 3.8-21.5 p=0.007                      | Commenting 13.7% less than Red Dot pre, 10.8% less post |
| 2 Effect          | Australia Neep RCT      | Baseline                        | Intensive 13.5h 2-day > non-intensive weekly                                                  | 1-week p=0.002, 12-week p<0.001, confidence improved    |
| 2 Effect          | UK Hargreaves online    | Sens 78.9% spec 79.0% acc 78.9% | Sens 78.8% spec 89.9% acc 86.0%                                                               | Spec +10.9% pneumothorax > skeletal                     |
| 2 Effect          | Singapore Del Gante     | -                               | Sens 87.4-98.9% mean 94.3% spec 96.4-99.9% mean 98.2% acc 93.1-99.5% mean 96.9% 16,483 images | High performance blended                                |
| 2 Sustainability  | Brown 2019 Australia    | 92.0% acc 71.1% sens 98.4% spec | Sustained 12 months 6,290 exams 100/month audit                                               | Consistent high                                         |
| 3 Reporting level | Piper 2005 UK           | -                               | >95% sens/spec/acc 6,796 OSE 28 radiographers                                                 | Postgraduate certificate                                |
| 3 Reporting level | Piper 2014 UK chest     | -                               | Sens 95.4% CI 94.4-96.3% agreement 89% 4,000 reports 40 radiographers                         | Masters-level                                           |
| 3 Reporting level | Woznitza 2018 UK JAFROC | -                               | FOM 0.828 vs 0.788 radiologists                                                               | p<0.0001 non-inferiority                                |

**The post training indicated significant positive effect compared to the baseline performance.**

**Table 6: GRADE Summary of Findings and Certainty**

| Outcome             | No. Studies | Effect Range                                                | Certainty | Reasons                                                                                                                                                                        |
|---------------------|-------------|-------------------------------------------------------------|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Diagnostic accuracy | 15 pre-post | 68.8%→83.3% +14.5% p=0.017, 71.7%→79.5% p=0.002, 71.04%→78% | Low       | Downgrade: serious RoB (4 serious), serious inconsistency (68.8-96.9% heterogeneous banks/ref), moderate indirectness (UK/Australia postgrad not applicable Nigeria), moderate |

|                                                                                |               |                                                                                                                                                                                    |          |                                                                                                                                                                                                   |
|--------------------------------------------------------------------------------|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                                |               | p=0.008, 89.9%→95%,<br>77.55%→80.87%                                                                                                                                               |          | imprecision (small n Africa 8-36),<br>suspected publication bias; Upgrade:<br>large effect +14.5% Ghana, dose-<br>response intensive > non-intensive<br>p<0.001, feedback sustains<br>91.9%→96.5% |
| Sensitivity                                                                    | 18<br>studies | 69.2%→83.3%<br>p=0.042,<br>84.1%→90.4%<br>p=0.012,<br>83.73%→87.28%,<br>0.71→0.81 Red Dot,<br>71.1% sustained,<br>94.3% mean Singapore,<br>95.4% Piper chest,<br>79.1% Takapautolo | Low      | Same as accuracy                                                                                                                                                                                  |
| Specificity                                                                    | 18<br>studies | 68.3%→83.3%<br>p=0.011, 80%→85%<br>p=0.176 NS,<br>59.62%→70.34%,<br>~0.95-0.96 maintained,<br>98.4% sustained,<br>98.2% mean, 95.8%<br>Takapautolo                                 | Low      | Same as accuracy                                                                                                                                                                                  |
| Descriptive<br>skills<br>(localisation,<br>description,<br>comment<br>quality) | 5 studies     | Loc 59.2%→68.4%<br>p=0.003, desc<br>55.3%→77.1%<br>p<0.001, diag<br>43.1%→49.3%<br>p=0.022, incorrect<br>comments<br>24.1%→17.78%<br>p=0.05                                        | Moderate | Downgrade moderate RoB, upgrade<br>large effect description +21.8%,<br>consistent improvement, low<br>inconsistency                                                                               |

GRADE certainty was Low for diagnostic accuracy (15 studies), sensitivity (18) and specificity (18) due to serious risk of bias, inconsistency (68.8-96.9% range), indirectness, imprecision (small African n=8-36) and

publication bias, but upgraded for large effect +14.5% and dose-response intensive better p<0.001; Moderate for localisation/description (5 studies) due to large effect +21.8% and consistency.