Lockwood, P. and Pittock, L. (2019) Multi-professional image interpretation: performance in preliminary clinical evaluation of appendicular radiographs. Radiography. pp. 1-13. ISSN 1078-8174.

Link to official URL (if available):

https://doi.org/10.1016/j.radi.2019.04.013

Multi-professional image interpretation: performance in preliminary clinical evaluation of appendicular radiographs

Lockwood P, Pittock L.

Clinical and Medical Sciences Research Hub, School of Allied & Public Health Professions, Canterbury Christ Church University, Kent, UK

Introduction

Within the United Kingdom (UK) various approaches to radiographic abnormality detection systems (RADS) have been in clinical practice since the start of the 20th century.[1] The current preliminary clinical evaluation (PCE) guidance formulated by the Society and College of Radiographers[2] (SCoR) is defined as a commentary on imaging findings to referrers. The application of PCE is not intended as a substitute for a definitive report but as a form of RADS. The objective of PCE is to support the referrer's clinical judgements[3-6] to aid initial patient treatment and management for acute injuries.

The Health and Care Professions Council (HCPC) Standards of proficiency: Radiographers[7] requires radiographers to "be able to distinguish between normal and abnormal appearances evident on images (13.14[7]), understand the structure and function of the human body in health, disease and trauma, as well as common pathologies and mechanisms of disease and trauma (13.21[7]), be able to formulate specific and appropriate management plans (14.2[7]) be able to undertake and record a thorough, sensitive and detailed clinical assessment" (14.6[7]). The SCoR[2] promotes the view that PCE is a core part of radiographers' scope of practice, and should be underpinned by education and competency development. The application of PCE into the scope of practice of other healthcare professions in the acute setting requires workforce development and professional body guidance documents specific to individual roles.

At present, studies[8-12] have shown the potential for radiographers to produce PCEs, with limited up to date literature demonstrating the image interpretation ability of other healthcare professionals[13-21] involved with the treatment and management of patients requiring radiographic imaging for trauma. The College of Paramedics[21] definition of the role of an advanced paramedic practitioner, in line with the Royal College of Emergency Medicine ACP pathway[22] for managing patients through accident and emergency (A&E) and urgent care pathways, includes requesting radiographic examinations and image interpretation. Similar to the role of emergency nurse practitioner (ENP) role that has been long established in the NHS.[23-27] Likewise, the multidisciplinary response to emergency department demands has also included the development of the role of emergency physiotherapists[28-33] to triage, manage and discharge patients, extending the scope of practice to include requesting and interpreting radiographic imaging. Drivers for PCE implementation include Department of Health (DoH) priorities[34,35] to improve healthcare efficiency in the National Healthcare Service (NHS) against workforce shortages and raising patient demand. In assessing the NHS, the Care Quality Commission[34,36] and the DoH[37] have advised that to continue delivering sustainable healthcare delivery the NHS must support local improvements in quality of care and patient outcomes.

Image interpretation of radiographs for healthcare professionals is now beginning to be embedded by UK universities as part of the multidisciplinary clinical assessment skills of Advanced Clinical Practitioners (ACP)[20] working in the acute setting. The short course at (*Anonymised for peer review* *manuscript)* University is a part-time interprofessional healthcare non-credit bearing short course for practitioners who interpret radiographic findings, influencing immediate patient management. The learning and teaching provided the skills necessary to produce PCE commentaries for trauma imaging of the appendicular skeleton. The campus-based study included image interpretation workshops (12 sessions) covering radiographic anatomy, pathology, osteology, normal variants, and search strategies with lectures on descriptive terminology, common errors, auditing, legal and professional issues by academic staff, reporting radiographers and consultant radiologist and acute medicine physician clinical practice supervisors. Supplementary sessions were provided for participants without a radiography background to introduce principles of radiographic techniques to recognise poor technical adequacy (positioning and exposure) which may impede interpretation. Educational support was provided by web-based radiographic image banks, on-line anatomy and pathology resources for off-campus viewing and clinical practice learning agreements to gain experience and exposure to prospective workloads. The assessment involved formative PCE commentaries completed in the participant's clinical workplace and a summative Objective Structured Examination (OSE) on campus.

Aims

This paper evaluates the educational programme OSE results of three cohorts of multidisciplinary healthcare professionals' interpretation of appendicular radiographic examinations at the end of the PCE module. The null hypothesis proposes no difference exists in the means of the test scores in the set of given observations, regardless of the professional backgrounds (variables of previous experience and training) as all the participants attained the same learning and teaching. The opposite may be shown through the observer performance that there is a difference in the means of the test scores (alternative hypothesis).

Methods

Approval for this study was received from the institutional ethics and governance approval panel. A sample size calculation was not performed for recruitment of participants in the study design (to gain a power (1-β) and effect size) as this study reports the results of students enrolled on a programme of study (fixed sample size).

Literature review of established standards

A literature search was completed to establish standards for comparison of the results of each subgroup (profession) in image interpretation of appendicular radiographic examinations (index test) of traumatic injuries (target conditions). The selection criteria of search terms included the observer group 'radiographer, nurse, paramedic, physiotherapist' and the index test used 'image interpretation, red dot, preliminary clinical evaluation' with Boolean operator search terms. Databases applied included Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, CINAHL, PubMed Central and Google Scholar, filtering results to the period of 1985 to 2018. Additional studies were sourced through reference lists in found papers. Found studies were reviewed for methodological quality against the QUADAS-2[38] criteria. In this study the published data needed to be identifiable for inclusion (results as true positive (TP), true negative (TN), false positive (FP), false negative (FN) or sensitivity, specificity, agreement) or excluded if data were incomplete or missing. Synthesis of available study data was performed using Meta-DiSc software[39] and displayed in forest plots displaying pooled estimates for sensitivity, and specificity analysis. Additional statistical meta-analysis calculated Chi-square (x^2) with p-values (p), and inconsistency I-square (I^2) to disclose any heterogeneity or consistency in the results, with degrees of freedom (*df*). Pooled log diagnostic odds ratio (DOR) was applied for any bias of study size and direction of effects. Final output measure calculated Summary Receiver Operating Characteristic (SROC) plots for each eligible subgroup result.

Radiographic test bank

The OSE caseload consisted of 25 retrospectively collected and anonymised appendicular skeletal radiographs (index test), with individual referral details of each case (9 females, age range 7-60 years; and 16 males, age range 4-57 years), presenting clinical information and referral source (12 from A&E, and 13 from general practice (GP). The abnormality prevalence consisted of 48% normal (including normal variant) cases and 52% abnormal cases (proportionate selection).[40-42] Internal validity of the cases ensured a range of subtle and textbook traumatic injuries (target conditions) were applied (Table 1) to reflect clinical practice.[43]

OSE Reference standard

The reference standard applied independent reporting (and blinded to the original radiologist report) of each OSE case by two reporting radiographers. The reports were reviewed for concordance of definitive answers (internal validity)[41,43] and range of target conditions to reflect the level and ability of skills of the module learning outcomes by the programme team.

Data collection

The OSE was undertaken in an academic environment under controlled conditions[40] using low-level lighting. Each radiographic case was displayed in Digital Imaging and Communications in Medicine (DICOM) format on KPACS software[44] to allow full image adjustment (windowing, zoom, pan, scroll, etc. of the data sets) by the participants.[45] The cases were displayed on DICOM compliant monitors (EIZO RadiForce RX350, 3MP, 54.1cm, 1536 x 2048 native resolution, 1,000 cd/m2 luminance, 1500:1 contrast ratio, DICOM Part 14 calibrated, with an anti-glare and anti-reflection coating) that complied with RCR image interpretation standards.[45]

The OSE instructions briefed the candidates to provide a confidence level score for each case as normal or abnormal and provide a detailed free text description of findings. In abnormal cases, the exact location (side and anatomical site), primary condition (single or multiple) and any secondary complications (swelling, angulation, dislocation, etc.) were required.

Test bank scores

The OSE scoring system applied a maximum two marks for abnormal (13 cases) and one mark for normal (12 cases) if correctly classified (including location and description), a total of 38 marks from the 25 cases was available. Fractionated scores were used for cases of more than one abnormality present.

Statistical analysis

The programme team scored the candidate answer booklets to the reference standard answers. Correct answers were graded as TP or TN; incorrect responses were classified as FP or FN. Fractions (whole and partial) were applied in cases of multiple target conditions as defined in a previous study[46] to allow calculation for sensitivity and specificity.[40] For individual reader performance Cohen's Kappa (reader performance reliability) was applied for inter-reader agreement with 95% confidence intervals (95% CI).[47] Subgroup (profession) analysis of OSE score, sensitivity, specificity with Analysis of Variance (ANOVA) for variation among and between the group's performance was applied (*alpha* = 0.05). Alternative Free-response Receiver Operating Characteristic (AFROC) methodology[48] was applied allowing for multiple pathological conditions and locations per case and Area Under the Curve (AUC) values estimated.

Results

Observer groups

The participants included nineteen nurses (N1-19; from ENP, GP and orthopaedic roles), and fourteen allied health professionals, seven radiographers (R1-7), five paramedics (P1-5) and two physiotherapists (PT1-2); Table 2. The demographics included twenty three females (age range 28-60 years, mean 40.4 years) and ten males (ages 28-48 years, mean 41 years). The clinical departments of the participants included four acute NHS district general hospitals and four minor injury NHS health centres with x-ray facilities.

Individual reader performance

The participant results demonstrated a range of individual test scores, sensitivity, specificity and Kappa outcomes (Table 2a-d). Participant group sizes varied which affected the range of scores for predication of overall mean analysis.

Subgroup test scores

The Kappa coefficient statistical measure of inter-rater agreement for the professional groups displayed a strong level of agreement for the physiotherapists (*k*=82), a good measure for the radiographers (*k*=72) and paramedics (*k*=61), and a moderate level of agreement for the nurses (*k*=56, Table 2a-d). The AFROC subgroup performance (Figure 1) displayed a limited range of confidence index ratings resulting in a constrained curve plotting of operating interval points. The subgroup AUC values were encouraging, with physiotherapists AUC 91.8 (95%CI 83-100), radiographers AUC 86.5 (95%CI 80.6-92.3); paramedics AUC 81.5 (95%CI 73.6-89.4), and nurses AUC 78.3 (95%CI 74-82.6).

The ANOVA estimated the variation (hypothesis testing) among and between the professional subgroups scores to calculate the differences among the sample means (table 3a-c). In this sample the exploratory data analysis result cannot reject the null hypothesis as ANOVA probability (*p*-value) for the OSE test scores (*p*=0.23) and OSE sensitivity score (*p*=0.65) were higher than the significance threshold (*alpha*=0.05), the OSE specificity score were equal but not lower (*p*=0.05) with the result not statically significant. To confirm this outcome the observed ANOVA values of *F* (Fisher–Snedecor distribution) against the *F* critical values (*F*-Crit), determined the *F* were less than the *F*-Crit (2.9) for the OSE test score, sensitivity, and specificity in all subgroup samples. A post-hoc power analyses on the sample size (to gain a power and effect size) would in effect repeat the *p*-value (ANOVA *F*-test confirmed result), in this situation the width and range of the 95% CI more appropriately reflects the statistical power (Tables 2a-d and Figure 1).

Subgroup performance comparison to literature

Comparison to published literature indicated a paucity of evidence from the subgroups which resulted in a wider literature search to include additional databases to confirm the limited pool of available studies for the nursing, paramedic and physiotherapist professions and provide a comprehensive review for scrutiny. A summary of the process and resulting studies are displayed in Figures 2, 3 and 4.

Synthesis of data included large-scale studies[12,49-53] that involved multiple rounds of assessment (pre, mid, post, and follow-up data)[9,11,12,16,53-62] which provided a wealth of quantitative results and relevant statistical information of performance ability. As such some studies were broken down into multiple entries (pre, mid, post, follow-up data)[9,11,12,16,53-62] for the pooled assessment per professional group. The highest amount of eligible studies for inclusion on reader performance in musculoskeletal radiographic image interpretation was found in the radiography profession (41 studies from 1985-2018[9,11,12,16,17,49-60,63-85]). Comparison of the pooled literature sensitivity 84% (95% CI 83-84%; Figure 3A) to this study 79.7% (95% CI 71.8-87.5%; Table 2a), pooled literature specificity 84% (95% CI 84-84%; Figure 3B) compared to this study 92.9% (95% CI 90.6-95.2%; Table 2a) and literature SROC AUC of 93.1 (Figure 3C) compared to the AFROC AUC 86.5 (Figure 1) in this study were encouraging. The pooled literature DOR (45.6; 95% CI 32.5-64.0; x^2 3158.9; *df* 60, *p*=0.00; I^2 98.1%) for the found studies displayed no concerning variation of DOR to represent publication bias.

The nursing profession evidence resulted in a small collection of eligible studies (from 1996-2014 [16,17,61,62,86-88]) which allowed extrapolation of results for analysis and construction of 2x2 contingency tables. Review of the pooled literature sensitivity 75% (95% CI 73-77%; Figure 4A) to this study 76.2% (95%CI 68-84.5%; Table 2b), and pooled literature specificity 80% (95%CI 77-82% Figure 4B) to the 80.4% (95%CI 75.1-85.8%; Table 2b) outcome in our study displayed strong results in this cohort performance. The literature SROC AUC of 91.2 (Figure 4C) in contrast to our study AFROC AUC of 78.3 (Figure 1) displayed minor variance. The results displayed in this study were equivalent to the found literature pooled results, the summary DOR (27.9; 95% CI 6.8-114.8; x^2 348.8; *df* 8, *p*=0.00; I^2 97.7%) for the literature did not suggest publication bias to be present.

Literature searching for skeletal radiographic image interpretation studies providing evidence for paramedic and physiotherapists abilities, identified no papers, thus limiting the comparison of these results to any standards for those professions.

Discussion

The participant results in this study were encouraging when compared to published studies assessing equivalent radiographic examinations[9,11,12,16,17,49-88] and applying similar scoring systems.[13,16,17] Radiographer image interpretation has been demonstrated extensively since the early work of Berman et al [49] and the systematic review by Brealey et al.[89] through to the work of Piper and Paterson[16] and Coleman and Piper[17] addressing a multidisciplinary approach and in recent studies of PCE ability of radiographers in the UK,[8] South Africa[9] and Australia.[11,59,81,85]

The literature review found a disposition towards qualitative research from the nursing profession discussing the role of requesting and triage of radiographic imaging in the acute setting from the UK[90-93] and Australia.[94,95] The nursing studies identified echoed the concerns raised by Snaith and

Hardy[62] of an absence of details to assess study validity, highlighting a lack of radiology reporting (radiologist or reporting radiographer) participation and instead applying inadequate reference standards (image interpretations) of junior doctors, details of assessment methods, characteristics of cases, and omitted full disclosure of results and data.[13,15,91,93,96]

Likewise, physiotherapists working in extended scopes of practice within the A&E setting to triage musculoskeletal injuries and request and interpret radiographs is a developing international practice (United States of America[97], Australia[98-101], UK[102,103], Ireland[104], Canada[105]). But a similar theme was noted throughout these found studies that there was a lack of radiology reporting (radiologist or reporting radiographer) involvement in either training or assessing physiotherapists in radiographic image interpretation and diagnosis (or applied as an appropriate reference standard/ground truth in observer performance studies). Kersten et al[106] explored the range of extended scope of practice of physiotherapists in the acute setting and raised similar concerns that due to the lack of professional guidance on radiographic image interpretation. Kersten et al[106] advise the physiotherapy professional body may wish to consider prerequisites of formal training before participating in this extended role to ensure the quality of care for patients, regulation and protection for the practitioners. These concerns apply to all professions, and any healthcare professional working at an advanced level practice should be supported by professional frameworks and training for safe service provision. However, Kersten et al[106] stopped short of explicitly highlighting radiology reporting (radiologist or reporting radiographer) input, or used as reference standards in training. More importantly, these studies fail to address the lack of rigorously applied methods to assess the ability of image interpretation[40] which should be adequately evaluated before implementation of the role extension in clinical practice. These problems were also reflected in the paramedic[107,108] literature.

Many variables potentially influence the results of the healthcare professionals enrolled in this study. The pre-existing skills and experience in physical clinical assessment of patients by nurses, paramedics and physiotherapists and utilisation of this knowledge can complement and enhance their clinical decision making in radiographic interpretation. Potentially this may provide an advantage over radiographers, although the undergraduate training, knowledge and experience at looking at radiographs, and the daily exposure and experience of this role in clinical practice by radiographers may likewise advantage them in the image interpretation task.

In examining the performance of the participants, errors occurred in the identification of upper limb cases containing multiple conditions. An example being an adult wrist with a fractured scaphoid (all noted) with an additional perilunate dislocation (missed by 16/33 participants, case 2). Likewise a paediatric elbow with raised fat pad and fracture of the medial epicondyle of the humerus (all noted) but with an additional fracture of the radial head (missed by 11/33, case 11). Adult shoulders also appeared in the FP errors with two normal cases having FP errors (6/33 case 4, 7/13 case7), and the posterior gleno-humeral joint dislocation was incorrectly called anterior by 9/33 (case 9). Of the lower limb cases, the paediatric ankle case raised the most FN errors for missed avulsion fractures (8/33; case 8) and FP errors for a normal paediatric ankle (9/33; case 15).

Training of the autonomous emergency practitioner and ACP roles are establishing a generic collection of learning skills in clinical and diagnostic assessment, decision making in management and treatment plans, pharmacology and discharge. The inclusion of radiographic image interpretation and PCE commentary training into ACP courses and CPD training for emergency healthcare professionals should attempt to train and assess individuals to an established benchmark score for safe clinical practice. Further research is recommended to investigate the utility of multidisciplinary PCE and image interpretation commenting downstream on decision making, time to treat, cost-effectiveness and patient outcomes.

Limitations

The cohort sample was small, and the particular professions were represented by inconsistent sample sizes (restricted paired data sets, producing inherent external validity bias) thus limiting the results outside of this study, a powered sample size is recommended for future studies. The literature data was not originally published in ROC methodology; thus we are aware caution should be applied comparing SROC AUC to AFROC AUC due to differences in calculations.

Conclusion

The comparison of image interpretation of appendicular radiographs in an academic environment by the radiographers and nurses to the published literature were encouraging (SROC AUC 93.1 and 91.2 respectively). The paramedics and physiotherapists results could not be compared to published standards or abilities on the task due to a lack of literature available. These small sample results, therefore, can provide an estimated baseline for future research. Radiographic image interpretation is now beginning to be embedded in multi-disciplinary ACP roles as an expectation by professional bodies, although with an absence of established standards for some healthcare professions to achieve in this task it raises potential regulatory, professional, and organisational safe practice concerns.

Conflict of interest statement

No conflict of interest, financial or otherwise, to declare.

Acknowledgements

We would like to thank all the healthcare professionals who participated in this study. The work described was carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans.

References

1. Moody, I. 50 years of History. London: The Society of Radiographers; 1970.
2. Society and College of Radiographers. Preliminary Clinical Evaluation and Clinical Reporting by Radiographers: Policy and Practice Guidance. London: Society and College of Radiographers; 2013.
3. Woznitza N, Piper K, Rowe S, West C. Optimizing patient care in radiology through team-working: a case study from the United Kingdom. Radiography 2014; 20(3):258-63.
4. Snaith B, Hardy M. Radiographer abnormality detection schemes in the trauma environment—An assessment of current practice. Radiography 2008; 14(4):277-81.
5. Snaith B, Hardy M, Lewis EF. Reducing image interpretation errors–Do communication strategies undermine this?. Radiography 2014; 20(3):230-4.

6. Lancaster A, Hardy M. An investigation into the opportunities and barriers to participation in a radiographer comment scheme, in a multi-centre NHS trust. Radiography 2012; 18(2):105-8.
7. The Health and Care Professions Council. Standards of Proficiency: Radiographers. London: HCPC; 2013.
8. Stevens BJ, Thompson JD. The impact of focused training on abnormality detection and provision of accurate preliminary clinical evaluation in newly qualified radiographers. Radiography. 2018 Feb 1;24(1):47-51.
9. Hazell L, Motto J, Chipeya L. The Influence of Image Interpretation Training on the Accuracy of Abnormality Detection and Written Comments on Musculoskeletal Radiographs by South African Radiographers, Journal of Medical Imaging and Radiation Sciences 2015, 46, 3, (302).
10. McConnell J, Devaney C, Gordon M. Queensland radiographer clinical descriptions of adult appendicular musculo-skeletal trauma following a condensed education programme, Radiography 2012, 19, pp. 48-55.
11. McConnell J, Devaney C, Gordon M, Goodwin M, Strahan R, Baird M. The impact of a pilot education programme on Queensland radiographer abnormality description of adult appendicular musculo-skeletal trauma. Radiography 2013, 18, pp. 184-190.
12. Hardy M, Culpan G. Accident and emergency radiography: a comparison of radiographer commenting and 'red dotting'. Radiography 2007, 13, pp. 65-71.
13. Meek S, Kendall J, Porter J, Freij R. Can accident and emergency nurse practitioners interpret radiographs? A multicentre study. J Accid Emerg Med 1998, 15, pp. 105-107.
14. Crane J, Delany C. Physiotherapists in emergency departments: responsibilities, accountability and education. Physiotherapy 2013, 99, pp. 95-100.
15. Overton-Brown P,Anthony D. Towards a partnership in care: nurses' and doctors' interpretation of extremity trauma radiology. J Adv Nurs, 27 (1998), pp. 890-896.
16. Piper KJ, Paterson A. Initial image interpretation of appendicular skeletal radiographs: a comparison between nurses and radiographers. Radiography 2009, 15, pp. 40-48.
17. Coleman L, Piper K. Radiographic interpretation of the appendicular skeleton: a comparison between casualty officers, nurse practitioners and radiographers. Radiography, 15 (2009), pp. 196-202.
18. Mordy M, Barrett C. Interpreting Trauma Radiographs Journal of Advanced Nursing 2003, October, vol 44, 81-87.
19. Hardy M, Barret C. Interpretation of trauma radiographs by radiographers and nurse practitioners in the UK: a comparative study. The British Journal of Radiology 2004, 77, 657-661.
20. Health Education England. Multi-professional framework for England. Health Education England 2017 Leeds.
21. The College of Paramedics and Health Education England. Digital Career Framework 2017. The College of Paramedics 2017, Bridgwater.
22. The Royal College of Emergency Medicine. Emergency Care Advanced Clinical Practitioner Curriculum and Assessment Adult Only/Adult and Paediatric / Paediatric. Version 2.0. The Royal College of Emergency Medicine 2017 London.
23. Chudleigh J. Nurse requested x –rays in A&E Department. Emergency Nurse 2004, vol 11, 32-35.
24. Ward W. Key Issues in Nurse requesting x-rays Emergency Nurse 1999. Feb, Vol 6, No 9.
25. Nursing and Midwifery Council. The Code: Standards of conduct performance and ethics for nurses and midwives (2008). London: Nursing and Midwifery Council.
26. Nursing and Midwifery Council. Record keeping: Guidance for nurses and midwives (2009) London: Nursing and Midwifery Council.
27. Making a Difference (1999). Strengthening the Nursing and Midwifery and Health Visiting Contribution to Health and health care. D.O.H London.

28. Chartered Society of Physiotherapy. Advanced practice in physiotherapy: Understanding the contribution of advanced practice in physiotherapy to transforming lives, maximising independence and empowering populations. Chartered Society of Physiotherapy 2016, London.
29. McPherson P, Kersten P, George S, Lattimer V, Ellis B, Breton A, et al. Extended roles for allied health professionals in the NHS: National Co-ordinating Centre for NHS Service Delivery and Organisation; 2004.
30. Taylor D, Bennett D, Cameron P. A paradigm shift in the nature of care provision in emergency departments. Emergency Medicine Journal 2004, 21, pp. 681-684.
31. Anaf S, Sheppard L. Describing physiotherapy interventions in the emergency department setting: an observational pilot study. Accid Emerg Nurs 2007, 15, pp. 34-39.
32. Kempson S. Physiotherapy in an accident and emergency department. Accid Emerg Nurs 1996, 4, pp. 198-202.
33. Anaf S, Sheppard LA. Physiotherapy as a clinical service in emergency departments: a narrative review. Physiotherapy. 2007 Dec 1;93(4):243-52.
34. The Care Quality Commission. The state of health care and adult social care in England 2016/17 (Presented to Parliament pursuant to section 83(4)(a) of the Health and Social Care Act 2008). Newcastle upon Tyne: House of Commons; 2017.
35. Department of Health. Improving outcomes: A Strategy for Cancer. London: HMSO; 2011.
36. The Care Quality Commission. Radiology review: A national review of radiology reporting within the NHS in England. The Care Quality Commission, Newcastle upon Tyne 2018
37. NHS England. Next steps on the NHS Five Year Forward View (Gateway number 06669). Leeds: NHS England; 2017.
38. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of internal medicine. 2011 Oct 18;155(8):529-36.
39. Zamora J, Abraira V, Muriel A, Khan KS, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Medical Research Methodology 2006, 6:31. Available from ftp://ftp.hrc.es/pub/programas/metadisc/Metadisc_update.htm 02.11.2016
40. Brealey S. Measuring the effects of image interpretation: an evaluative framework. Clinical Radiol 2001; 56.341-347.
41. Brealey S, Scally AJ. Methodological approaches to evaluating the practice of radiographers' interpretation of images, A Review. Radiography 2008, 14 (1): 46-54.
42. Piper K, Ryan C, Paterson A. The Implementation of a Radiographic Reporting Service, for trauma examinations of the skeletal system, in 4 National Health Service Trusts. Project Report SPGS 438. South Thames Regional Office (NHSE); 1999.
43. Robinson PJ, Wilson D, Coral A, Murphy A, Verow P. Variation between experienced observers in the interpretation of accident and emergency radiographs. The British Journal of Radiology 1999, 72(856): 323-330.
44. IMAGE Information Systems Ltd. K-PACS (Version 1.6.0). Germany, 2008.
45. Royal College of Radiologists. Picture archiving and communication systems (PACS) and guidelines on diagnostic display devices. 2012. London, UK.
46. (*Anonymised for peer review manuscript)*.
47. Kundel HL, Polansky M. Measurement of observer agreement. Radiology. 2003 Aug;228(2):303-8.
48. Thompson JD, Manning DJ, Hogg P. Analysing data from observer studies in medical imaging research: an introductory guide to free-response techniques. Radiography. 2014 Nov 1;20(4):295-9.
49. Berman L, de Lacey G, Twomey E, Twomey B, Welch T, Eban R. Reducing errors in the accident department: a simple method using radiographers. Br Med J (Clin Res Ed). 1985 Feb 9;290(6466):421-2.

50. Bowman S. Introducing an abnormality detection system by radiographers into an accident and emergency department: an investigation into radiographers' concerns about the introduction of such a system. Res Radiography. 1991;1:2-0.
51. Renwick IG, Butt WP, Steele B. How well can radiographers triage x ray films in accident and emergency departments?. BMJ. 1991 Mar 9;302(6776):568-9.
52. Loughran CF. Reporting of fracture radiographs by radiographers: the impact of a training programme. Br J Radiol 1994; 67: 945e50
53. Mackay SJ. The impact of a short course of study on the performance of radiographers when highlighting fractures on trauma radiographs: "The Red Dot System". The British journal of radiology. 2006 Jun;79(942):468-72.
54. Morrison ID, Hendry R, Fell C, Stothard K. An audit of radiographers' accuracy in recording abnormalities on causality radiographs using the red dot protocol. Clinical Radiology. 1999;12(54):862-3.
55. Ross C. The effectiveness of a short course of study on the ability of radiographers to highlight and describe abnormalities on skeletal accident and emergency films. MSc Thesis. South Bank University; 2002.
56. McConnell JR, Webster AJ. Improving radiographer highlighting of trauma films in the accident and emergency department with a short course of study--an evaluation. The British journal of radiology. 2000 Jun;73(870):608-12.
57. Hargreaves J, Mackay S. The accuracy of the red dot system: can it improve with training? Radiography (2003) 9, 283-289.
58. Hazell, L.J., 2012. A study to assess the ability of radiographers to apply pattern recognition criteria and interpret radiographs (Doctoral dissertation, University of Johannesburg) 2012.
59. Williams I, Baird M, Pearce B, Schneider M. Improvement of radiographer commenting accuracy of the appendicular skeleton following a short course in plain radiography image interpretation: A pilot study. Journal of medical radiation sciences. 2018 Oct 9.
60. Tay YX, Wright C. Image interpretation: Experiences from a Singapore in-house education program. Radiography. 2018 May 4.
61. Derksen RJ, Bakker FC, Heilbron EA, Geervliet PC, Spaans IM, de Lange-de Klerk ESM, Veenings B, Patka P, Henk JTM. Diagnostic accuracy of lower extremity X-ray interpretation by 'specialized' emergency nurses. Eur J Emerg Med, 13 (2006), pp. 3-8.
62. Snaith B, Hardy M. Emergency department image interpretation accuracy: The influence of immediate reporting by radiology. International emergency nursing. 2014 Apr 1;22(2):63-8.
63. Khiroya T, Taylor S, Remedios D, de Lacey G.Trauma radiology: the accuracy of radiographers' assessment of appendicular trauma Br J Radiol, 70 (Suppl) (1997), p. 123.
64. Orames C. Emergency department X-ray diagnosis-how do radiographers compare?. Radiographer: The Official Journal of the Australian Institute of Radiography, The. 1997 Apr;44(1):52-55.
65. Hall R, Egan I, Kleeman SJ. The red dot system: the outback experience. Radiographer: The Official Journal of the Australian Institute of Radiography, The. 1999 Aug;46(2):83.
66. Remedios D, Ridley N, Taylor S, de Lacey G. Trauma radiology: extending the Red Dot system. Br J Radiol 1998; 71(S):60.
67. Smith T, Younger C. Accident and emergency radiological interpretation using the radiographer opinion form (ROF). Radiographer 2002; 49: 27–31.
68. Rolph H. Red dot radiographers: assessment of their ability to detect abnormalities and write a short quality description. BSc Thesis. Glasgow Caledonian University, 2003.
69. Mc Entee MF, Dunnion S. A FROC analysis of radiographers performance in identification of distal radial fractures. European Journal of Radiography. 2009 Sep 1;1(3):90-4.
70. Rutledge A, McEntee MF, Rainford L, O'Grady M, McCarthy K, Butler ML. The impact of clinical indications on visual search behaviour in skeletal radiographs. InMedical Imaging

2011: Image Perception, Observer Performance, and Technology Assessment 2011 Mar 3 (Vol. 7966, p. 79660Y). International Society for Optics and Photonics.
71. Brown N, Leschke P. Evaluating the true clinical utility of the red dot system in radiograph interpretation. Journal of medical imaging and radiation oncology. 2012 Oct;56(5):510-3.
72. Hlongwane ST, Pitcher RD. Accuracy of after-hour 'red dot' trauma radiograph triage by radiographers in a South African regional hospital. SAMJ: South African Medical Journal. 2013 Jan;103(9):638-40.
73. Wright C. RadBench: benchmarking image interpretation performance. Liverpool: UKRC; 2012. June, http://shura.shu.ac.uk/id/eprint/8432.
74. Yallappa S, Clarke D, Minford JE. Accuracy of red dot in clinical practice. Clinical Radiology. 2012 Sep 1;67:S20.
75. du Plessis J, Pitcher R. Towards task shifting? A comparison of the accuracy of acute trauma-radiograph reporting by medical officers and senior radiographers in an African hospital. Pan African Medical Journal. 2015;21(1).
76. Kranz R, Cosson P. Anatomical and/or pathological predictors for the “incorrect” classification of red dot markers on wrist radiographs taken following trauma. The British journal of radiology. 2015 Jan 19;88(1046):20140503.
77. Higgins S, Wright C. Traffic light: an alternative approach to abnormality signalling. Liverpool: UKRC; 2016. https://core.ac.uk/display/77055324.
78. Neep, M. Pozzias, E. To alert or not to alert: the rise of the frontline radiographer. The Australian Society of Medical Imaging and Radiation Therapy's 11th National Conference, Canberra, Australia, 22-24 March 2016, Journal of Medical Radiation Sciences, March 2016 Volume 63, Issue S1.
79. Wright C, Reeves P. RadBench: Benchmarking image interpretation skills. Radiography. 2016 May 1;22(2):e131-6.
80. Akimoto T, Wright C, Reeves P. Image interpretation performance: a multicentre undergraduate study. Manchester: UKRC; 2017. https://doi.org/10.13140/RG.2.2.36820.40328.
81. Neep, M. Brown, C. Reaching for new heights: Logan Hospital radiographer commenting. The Australian Society of Medical Imaging and Radiation Therapy's 12th National Conference, Canberra, Australia, 24-26 March 2017, Journal of Medical Radiation Sciences, March 2017 Volume 64, Issue S1.
82. Rodriguez CM. A student radiographer’s accuracy in ‘red-dotting’ appendicular skeletal radiographs: Clinical Audit. 2017
83. Tay YX, Wright C. Development of Image Interpretation skills - a follow up study of Singapore diagnostic radiographers. 2017.
84. Wright C, Reeves P. Image interpretation performance: a longitudinal study from novice to professional. Radiography 2017;23(1):e1e7.
85. Davis, B. Development of a reliable and sustainable radiographer commenting audit. The Australian Society of Medical Imaging and Radiation Therapy's 13th National Conference, Canberra, Australia, 15–18 March 2018, Journal of Medical Radiation Sciences, March 2018 Volume 65, Issue S1.
86. Freij RM, Duffy T, Hackett D, Cunningham D, Fothergill J. Radiographic interpretation by nurse practitioners in a minor injuries unit. Emergency Medicine Journal. 1996 Jan 1;13(1):41-43.
87. Benger JR. Can nurses working in remote units accurately request and interpret radiographs? Emerg Med J 2002;19: 68–70.
88. Lee GA, Chou K, Jennings N, O’Reilly G, McKeown E, Bystrzycki A, Varma D. The accuracy of adult limb radiograph interpretation by emergency nurse practitioners: A prospective comparative study. International journal of nursing studies. 2014 Apr 1;51(4):549-54.

89. Brealey S, Scally A, Hahn S, Thomas N, Godfrey C, Crane S. Accuracy of radiographers red dot or triage of accident and emergency radiographs in clinical practice: a systematic review. Clinical radiology. 2006 Jul 1;61(7):604-15.
90. Swaby-Larsen D. X-ray interpretation by emergency nurse practitioners: Dorthe Swaby-Larsen and colleagues discuss an audit of the ability of emergency nurses to interpret X-rays accurately. Emergency Nurse. 2009 Oct 1;17(6):24-9.
91. Tachakra S, Mukherjee P, Smith C, Dutton D. Are accident and emergency consultants as accurate as consultant radiologists in interpreting plain skeletal radiographs taken at a minor injury unit? Eur J Emerg Med, 9 (2) (2002), pp. 131-134.
92. Hardy ML, Barrett C. Requesting and interpreting trauma radiographs: a role extension for accident & emergency nurses. Accident and emergency nursing. 2003 Oct 1;11(4):202-13.
93. Summers A. Can nurses interpret X-rays safely without formal tuition?. Accident and emergency nursing. 2005 Jul 1;13(3):162-6.
94. McConnell J, Baird M. The accuracy of adult limb radiograph interpretation by emergency nurse practitioners: A prospective comparative study. International journal of nursing studies. 2015 Jan 1;52(1):495-7.
95. Free B, Lee GA, Bystrzycki A. Literature review of studies on the effectiveness of nurses ability to order and interpret X-rays. Australasian Emergency Nursing Journal. 2009 Feb 1;12(1):8-15.
96. Sakr M, Angus J, Perrin J, Nixon C, Nichol J, Wardrope J. Care of minor injuries by emergency nurse practitioners or junior doctors: a randomised controlled trial. The Lancet. 1999 Oct 16;354(9187):1321-6.
97. de Gruchy A, Granger C, Gorelik A. Physical therapists as primary practitioners in the emergency department: six-month prospective practice analysis. Physical therapy. 2015 Sep 1;95(9):1207-16.
98. Gill SD, Stella J. Implementation and performance evaluation of an emergency department primary practitioner physiotherapy service for patients with musculoskeletal conditions. Emergency Medicine Australasia. 2013 Dec;25(6):558-64.
99. Sutton M, Govier A, Prince S, Morphett M. Primary-contact physiotherapists manage a minor trauma caseload in the emergency department without misdiagnoses or adverse events: an observational study. Journal of physiotherapy. 2015 Apr 1;61(2):77-80.
100. Kinsella R, Collins T, Shaw B, Sayer J, Cary B, Walby A, Cowan S. Management of patients brought in by ambulance to the emergency department: role of the Advanced Musculoskeletal Physiotherapist. Australian Health Review. 2018 Jun 26;42(3):309-15.
101. Stanhope J, Grimmer-Somers K, Milanese S, Kumar S, Morris J. Extended scope physiotherapy for orthopedic outpatients: an update systematic review of the literature. J Multidiscip Healthc . 2012;5:37–45.
102. Ball ST, Walton K, Hawes S. Do emergency department physiotherapy Practitioner's, emergency nurse practitioners and doctors investigate, treat and refer patients with closed musculoskeletal injuries differently?. Emergency Medicine Journal. 2007 Mar 1;24(3):185-8.
103. Daker-White G, Carr AJ, Harvey I, Woolhead G, Bannister G, Nelson I, Kammerling M. A randomised controlled trial. Shifting boundaries of doctors and physiotherapists in orthopaedic outpatient departments. Journal of epidemiology & community health. 1999 Oct 1;53(10):643-50.
104. Fennelly O, Blake C, FitzGerald O, Breen R, O'Sullivan C, O'Mir M, Desmeules F, Cunningham C. Advanced musculoskeletal physiotherapy practice in Ireland: A National Survey. Musculoskeletal care. 2018 Jun 21.
105. Desmeules F, Roy JS, MacDermid JC, Champagne F, Hinse O, Woodhouse LJ. Advanced practice physiotherapy in patients with musculoskeletal disorders: a systematic review. BMC Musculoskelet Disord . 2012;13:107.

106. Kersten P, McPherson K, Lattimer V, George S, Breton A, Ellis B. Physiotherapy extended scope of practice–who is doing what and why?. Physiotherapy. 2007 Dec 1;93(4):235-42.
107. Woollard M. The role of the paramedic practitioner in the UK. Australasian Journal of Paramedicine. 2015 Jul 15;4(1).
108. College of paramedics. Paramedic post-graduate curriculum guidance 2017. College of paramedics. The College of Paramedics 2017, Bridgwater.

Table 1. Test bank cases.

1.	Normal Paediatric Tibia and Fibula
2.	Abnormal Adult Hand [Fracture and Dislocation]
3.	Abnormal Adult Hand [Fracture]
4.	Normal Adult Shoulder
5.	Abnormal Adult Knee [Fracture]
6.	Abnormal Paediatric Tibia and Fibula [Fracture]
7.	Normal Adult Shoulder
8.	Abnormal Adult Ankle [Avulsion Fracture]
9.	Abnormal Adult Shoulder [Dislocation]
10.	Abnormal Adult Knee [Avulsion Fracture]
11.	Abnormal Adult Elbow [Multiple Fracture]
12.	Normal Adult Ankle
13.	Normal Adult Scaphoid
14.	Abnormal Paediatric Clavicle [Fracture]
15.	Normal Paediatric Ankle
16.	Normal Adult Fingers
17.	Normal Paediatric Elbow
18.	Abnormal Adult Foot [Multiple Fractures]
19.	Normal Paediatric Foot
20.	Abnormal Adult Foot [Fracture and Dislocation]
21.	Abnormal Paediatric Finger [Fracture]
22.	Normal Adult Elbow
23.	Normal Adult Wrist
24.	Normal Adult Knee
25.	Abnormal Paediatric Ankle [Fracture]

Table 2a. Radiographers R1-7 individual performance.

Participant	*Test Score (%)*	*Sensitivity (95%CI)*	*Specificity (95%CI)*	*Cohens Kappa (95%CI)*
R1	36.5 (96%)	92.3 (72.4-92.3)	100 (78.4-100)	92.0 (50.6-92.0)
R2	29.5 (78%)	65.4 (43.5-72.7)	91.7 (67.9-99.6)	56.4 (11.2-71.4)
R3	32.25 (85%)	78 (55.2-85.6)	92 (69.2-99.6)	70 (24.5-85.1)
R4	30 (79%)	67.3 (45.3-74.6)	91.7 (67.9-99.6)	58.4 (13.1-7.34)
R5	34.5 (90%)	88.5 (67-95.7)	91.7 (68.4-99.5)	80 (35.4-95.1)
R6	34.5 (90%)	88.5 (67-95.7)	91.7 (68.4-99.5)	80 (35.4-95.1)
R7	32.25 (85%)	78 (55.2-85.6)	92 (69.2-99.6)	70 (24.5-85.1)
Mean	32.7 (86%)	79.7 (71.8-87.5)	92.9 (90.6-95.1)	72.4 (63-81.7)
Median	32.2	78	91.7	70
Mode	32.2	78	91.7	70
Standard Deviation	2.5	10.6	3.1	12.6

Table 2b.Nurses N1-19 individual performance.

Participant	*Test Score (%)*	*Sensitivity (95%CI)*	*Specificity (95%CI)*	*Cohens Kappa (95%CI)*
N1	33 (87%)	98.1 (77.0-100)	75.0 (52.2-77.1)	73.7 (29.5-77.8)
N2	33 (87%)	98.1 (77.0-100)	75.0 (52.2-77.1)	73.7 (29.5-77.8)
N3	37 (97%)	100 (79.1-100)	94.1 (74.0-94.1)	94.0 (53.0-94.0)
N4	34 (88%)	100 (81.1-100)	72.7 (48.6-72.7)	74.9 (30.6-74.9)
N5	35.5 (93%)	86.5 (66.1-86.5)	100 (77.9-100)	86.1 (43.7-86.1)
N6	36.25 (95%)	98.2 (79.4-100)	90.9 (67.0-93.2)	89.8 (46.8-93.9)
N7	33.25 (88%)	91.4 (69.7-99.4)	83.3 (59.8-92.0)	74.9 (29.5-91.6)
N8	23.5 (62%)	59.1 (33.2-80.9)	64.3 (44-81.4)	23.3 (0-62.1)
N9	23 (61%)	58.1 (31.9-80.7)	63.2 (43.4-80.2)	21.1 (0-60.3)
N10	21 (55%)	59 (30.9-83.7)	52.5 (34.5-68.3)	10.8 (0-49)
N11	26 (69%)	62.7 (39.8-78.8)	75.5 (51.6-92.2)	38.1 (0-70.8)
N12	32.5 (86%)	87 (62.1-98.1)	85.2 (64-94.7)	71.9 (26-92.5)
N13	23 (60%)	42.9 (23.3-54.5)	81.8 (57-96.6)	23.3 (0-48.2)
N14	30.5 (80%)	76.9 (54.4-89.1)	83.3 (59-96.5)	60.1 (12.4-85.3)
N15	27 (71%)	61.8 (40.7-73.6)	82.2 (56.4-96.6)	42.9 (0-68.4)
N16	28.5 (75%)	66 (41.3-80.7)	83 (61.2-96.1)	49.4 (02.5-77.4)
N17	33 (87%)	81 (63.6-82.5)	97.3 (67.7-100)	73.7 (29.5-77.8)
N18	26 (68%)	57.1 (36.5-68.7)	81.8 (55.6-96.5)	37.5 (0-62.8)
N19	28 (73%)	65.7 (49-71.3)	87.9 (54-99.4)	46.7 (2.5-61.7)
Mean	29.6 (77.8%)	76.2 (67.9-84.4)	80.4 (75-85.7)	56.1 (44.6-67.5)
Median	30.5	76.9	82.2	60.1
Mode	33	98.1	75	73.7
Standard Deviation	4.9	18.3	11.9	25.5

Table 2c. Paramedics P1-5 individual performance.

Participant	Test Score (%)	Sensitivity (95%CI)	Specificity (95%CI)	Cohens Kappa (95%CI)
P1	29.75 (78%)	80.8 (58.2-94.4)	75.0 (50.6-89.8)	55.9 (08.8-84.3)
P2	32.25 (85%)	94.2 (72.6-99.9)	75.0 (51.6-81.1)	69.7 (24.4-81.6)
P3	29.75 (78%)	80.8 (58.2-94.4)	75.0 (50.6-89.8)	55.9 (08.8-84.3)
P4	30 (79%)	74.5 (49.6-88.8)	83 (61-95.8)	57.7 (10.7-84.9)
P5	32 (84%)	71.4 (51.9-71.4)	100 (75.2-100)	68.8 (26.1-68.8)
Mean	30.7 (80.7%)	80.3 (72.6-87.9)	81.6 (72.1-91)	61.6 (55.4-67.7)
Median	30	80.8	75	57.7
Mode	29.7	80.8	75	55.9
Standard Deviation	1.2	8.7	10.8	7.0

Table 2d. Physiotherapists PT1-2 individual performance.

Participant	Test Score (%)	Sensitivity (95%CI)	Specificity (95%CI)	Cohens Kappa (95%CI)
PT1	38 (100%)	100 (81.2-100)	100 (79.6-100)	100 (60.9-100)
PT2	31.25 (82%)	80.8 (58.4-92.7)	83.3 (59.1-96.3)	64.0 (17.4-88.9)
Mean	34.6 (91%)	90.4 (71.6-100)	91.6 (75.2-100)	82 (46.8-100)
Median	34.6	90.4	91.6	82
Mode	#N/A	#N/A	#N/A	#N/A
Standard Deviation	4.7	13.5	11.8	25.4

Table 3a. ANOVA of differences among group means in the study sample OSE Test score.

Groups	Count	Sum	Average	Variance
Radiographers	7	229.5	32.7	6.4
Nurses	19	564	29.6	24.7
Paramedics	5	153.7	30.7	1.5
Physiotherapists	2	69.2	34.6	22.7

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	P-value	F crit
Between Groups	80.5	3	26.8	1.5	0.23	2.9
Within Groups	512.9	29	17.6			
Total	593.4	32				

Table 3b. ANOVA of differences among group means in the study sample sensitivity score.

Groups	*Count*	*Sum*	*Average*	*Variance*		
Radiographers	7	558	79.7	112.9		
Nurses	19	1449.6	76.2	335.9		
Paramedics	5	401.7	80.3	76.6		
Physiotherapists	2	180.8	90.4	184.3		
Source of Variation	***Sum of Squares***	***Degrees of Freedom***	***Mean Square***	***F***	***P-value***	***F crit***
Between Groups	402.8	3	134.2	0.5	0.65	2.9
Within Groups	7215.2	29	248.8			
Total	7618	32				

Table 3c. ANOVA of differences among group means in the study sample specificity score.

Groups	*Count*	*Sum*	*Average*	*Variance*		
Radiographers	7	650.8	92.9	9.6		
Nurses	19	1529	80.4	141.8		
Paramedics	5	408	81.6	117.8		
Physiotherapists	2	183.3	91.6	139.4		
Source of Variation	***Sum of Squares***	***Degrees of Freedom***	***Mean Square***	***F***	***P-value***	***F crit***
Between Groups	945.4	3	315.1	2.8	0.05	2.9
Within Groups	3221.1	29	111			
Total	4166.6	32				

Figure 1. Alternative Free-response Receiver Operating Characteristic (AFROC) curve plot of professional subgroup performance.

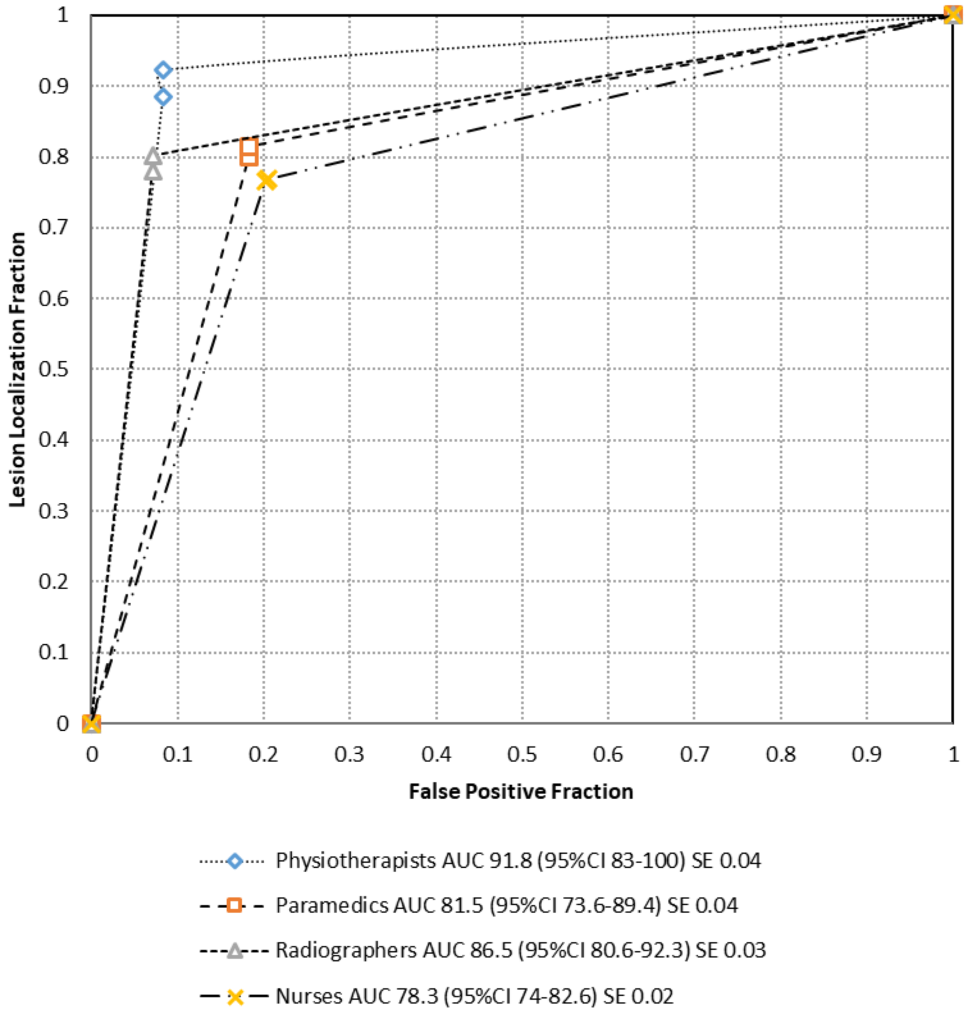

Figure 2. Literature search.
Figure 2. Literature search.

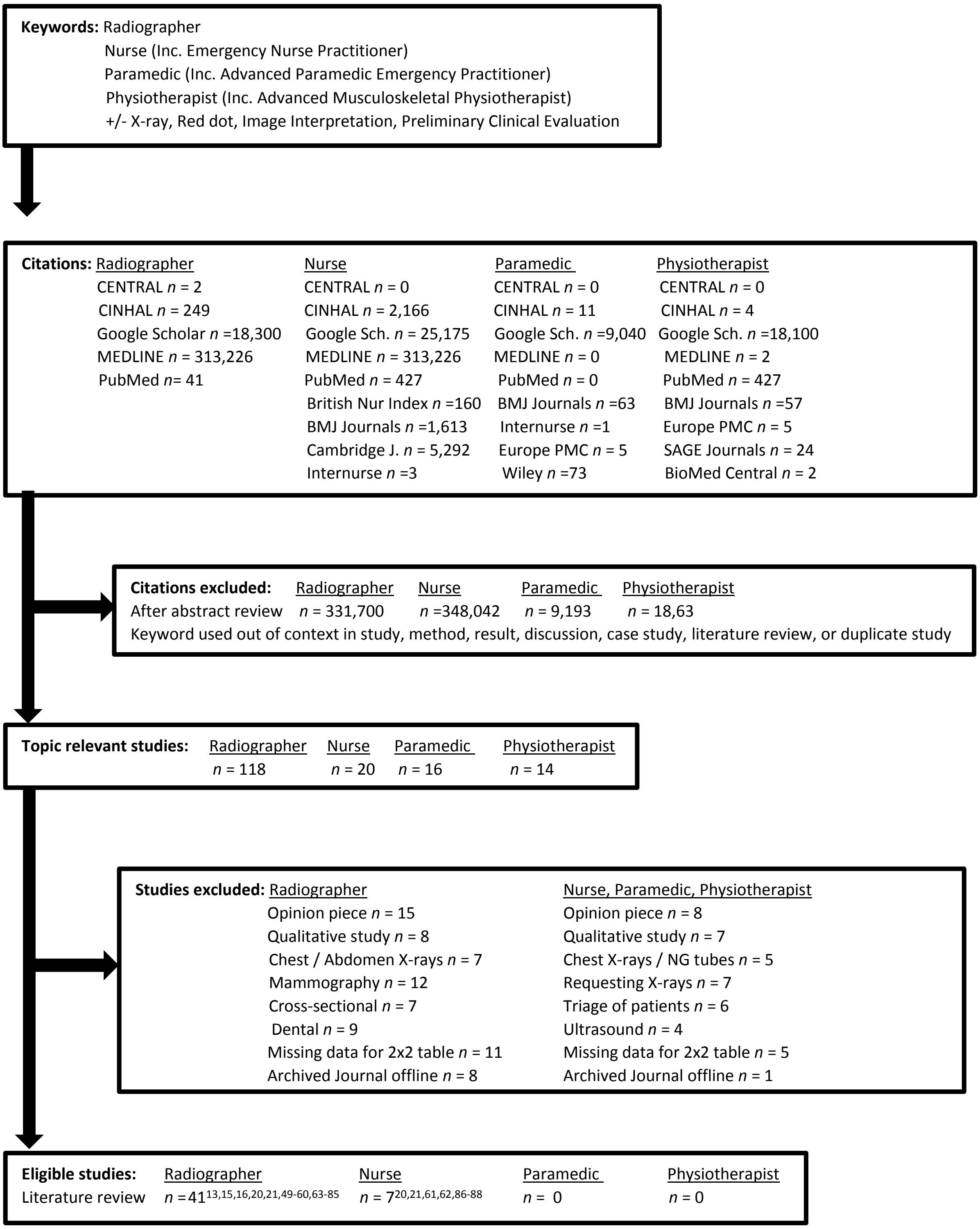

Figure 3 A. Radiographer literature analysis of sensitivity in musculoskeletal radiograph image interpretation (*pre-test results, “mid-test results, ~post-test results, ^ follow-up test results).

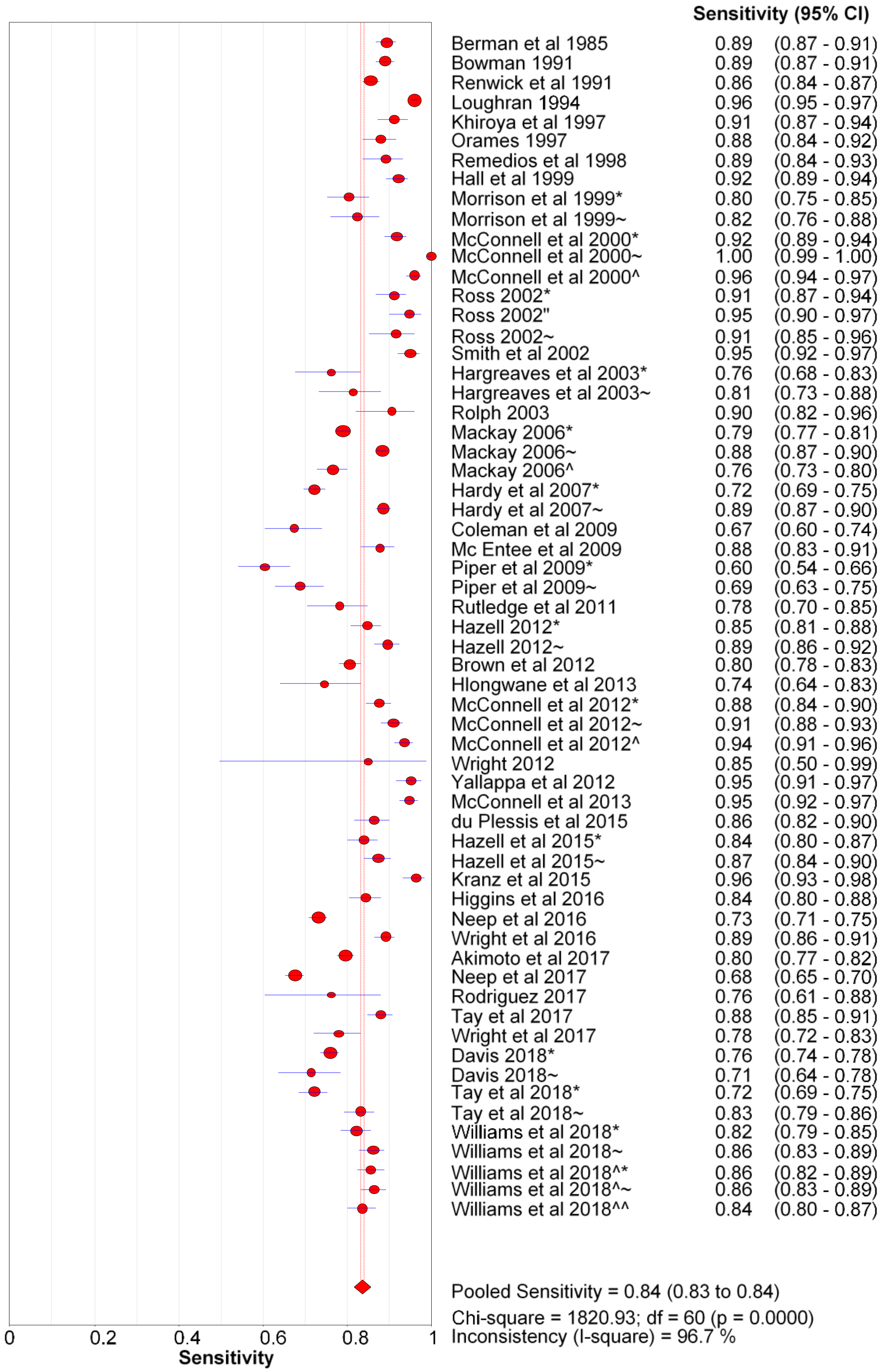

Figure 3 B. Radiographer literature analysis of specificity in musculoskeletal radiograph image interpretation (*pre-test results, “mid-test results, ~post-test results, ^ follow-up test results).

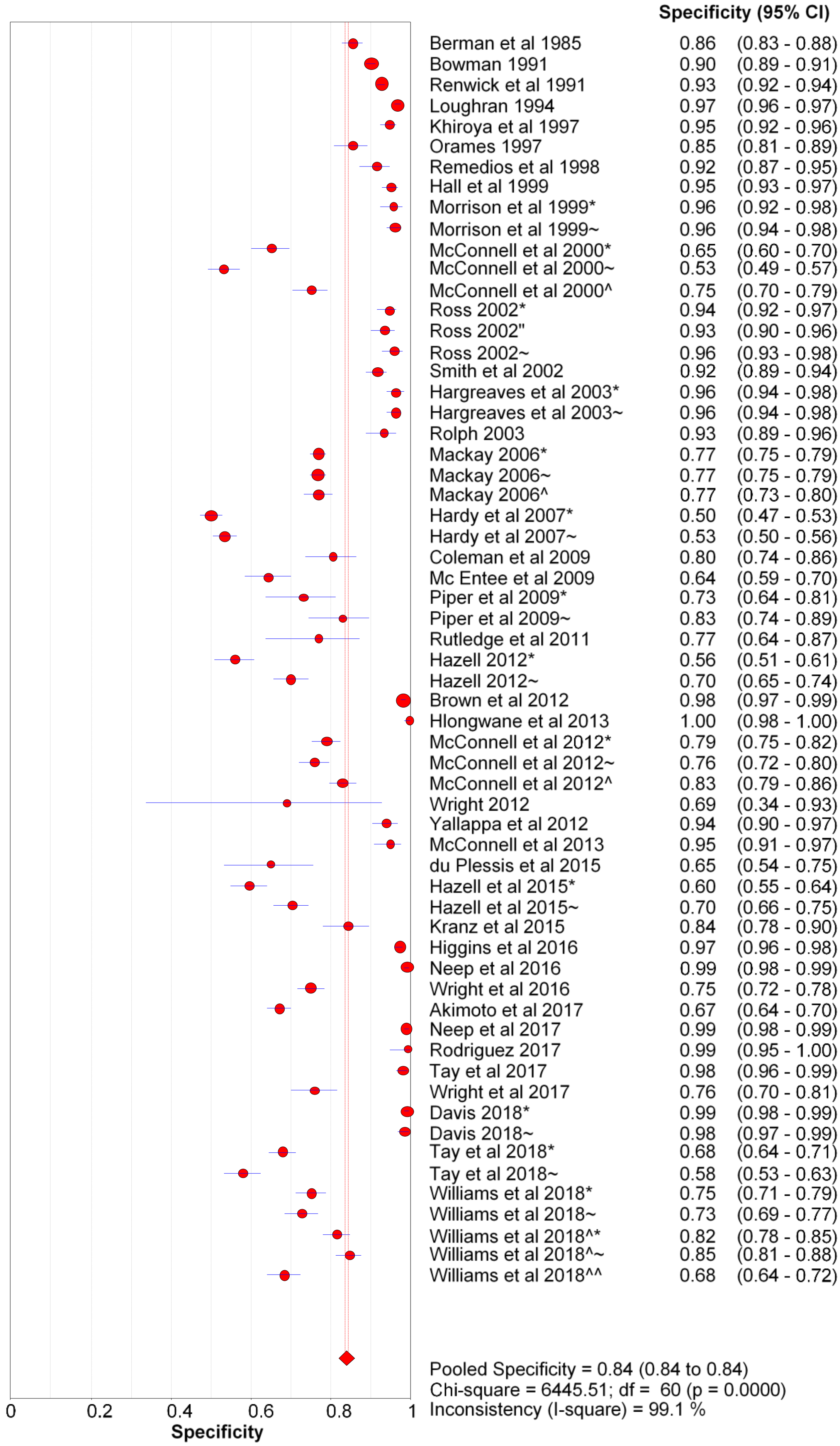

Figure 3 C. Radiographer Summary Receiver Operating Characteristic (SROC) curve analysis.

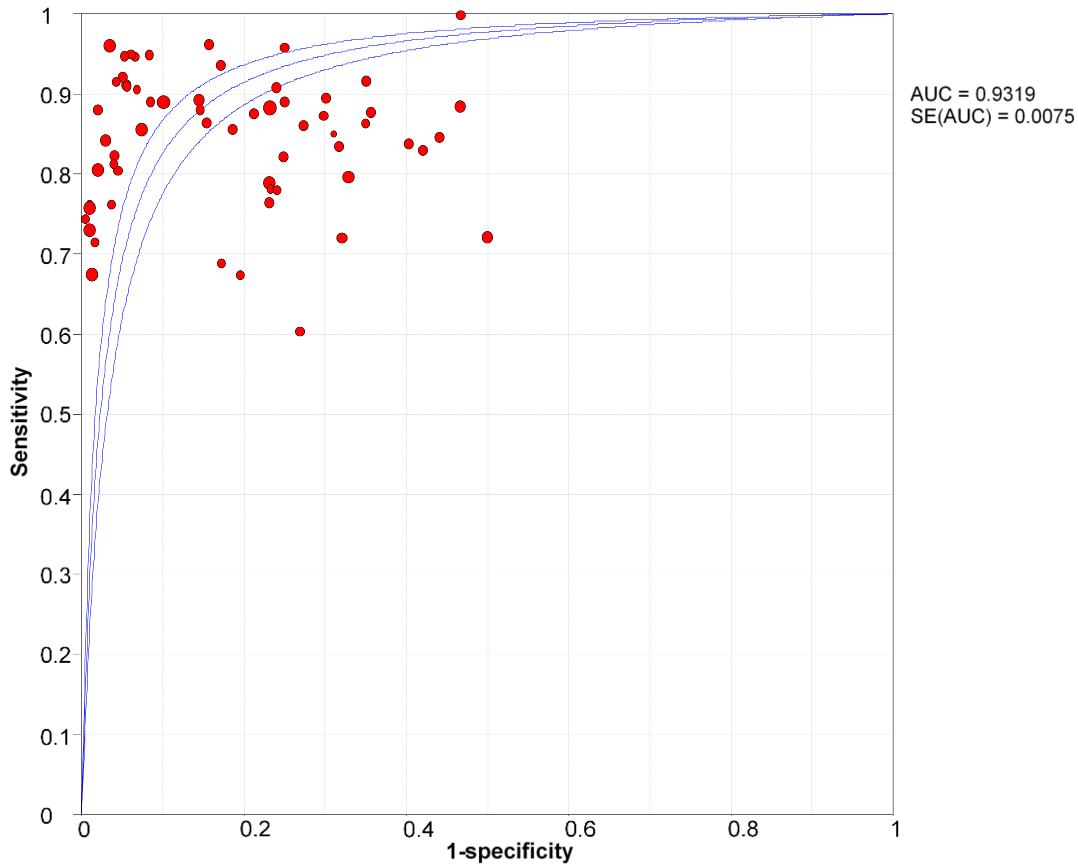

Figure 4 A. Nurse literature analysis of sensitivity in musculoskeletal radiograph image interpretation (*pre-test results, “mid-test results, ~post-test results, ^ follow-up test results).

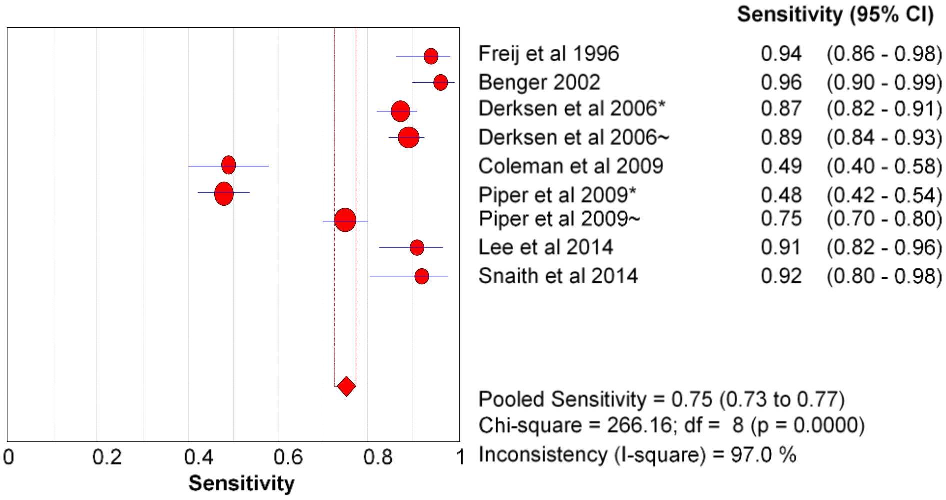

Figure 4 B. Nurse literature analysis of specificity in musculoskeletal radiograph image interpretation (*pre-test results, “mid-test results, ~post-test results, ^ follow-up test results).

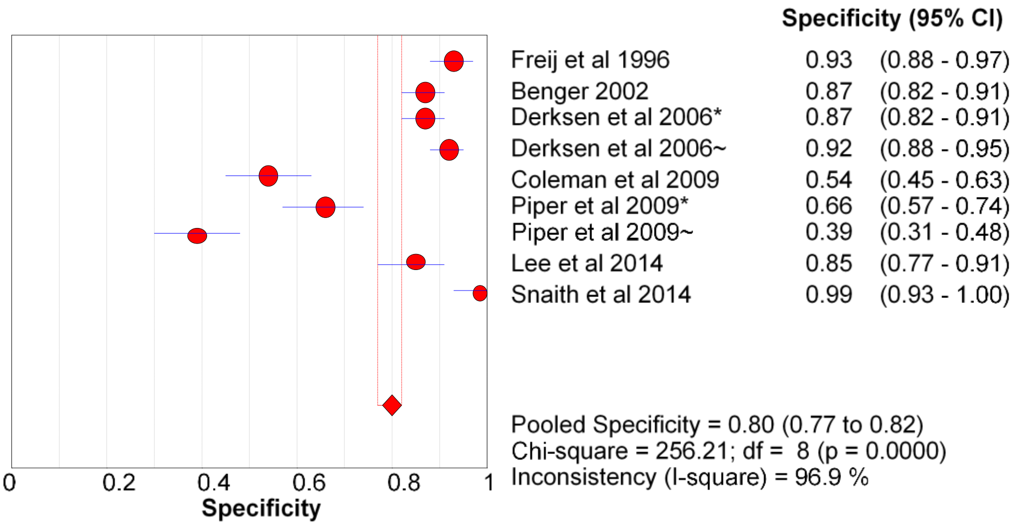

Figure 4 C. Nurse Summary Receiver Operating Characteristic (SROC) curve analysis.