Radiography xxx (2017) 1–12

Contents lists available at ScienceDirect

Radiography

journal homepage: www.elsevier.com/locate/radi

Review article

Computer-aided detection in musculoskeletal projection radiography: A systematic review

M. Gundry [a, *], K. Knapp [a], R. Meertens [a], J.R. Meakin [b]

[a] University of Exeter Medical School, St Luke's, Heavitree Road, Exeter EX1 2LU, UK
[b] College of Engineering Mathematics and Physical Sciences, University of Exeter, Stocker Road, Exeter EX4 4QL, UK

ARTICLE INFO

Article history:
Received 24 July 2017
Received in revised form
31 October 2017
Accepted 16 November 2017
Available online xxx

Keywords:
Computer-aided detection
CAD
MSK

ABSTRACT

Objectives: To investigated the accuracy of computer-aided detection (CAD) software in musculoskeletal projection radiography via a systematic review.
Key findings: Following selection screening, eligible studies were assessed for bias, and had their study characteristics extracted resulting in 22 studies being included. Of these 22 three studies had tested their CAD software in a clinical setting; the first study investigated vertebral fractures, reporting a sensitivity score of 69.3% with CAD, compared to 59.8% sensitivity without CAD. The second study tested dental caries diagnosis producing a sensitivity score of 68.8% and specificity of 94.1% with CAD, compared to sensitivity of 39.3% and specificity of 96.7% without CAD. The third indicated osteoporotic cases based on CAD, resulting in 100% sensitivity and 81.3% specificity.
Conclusion: The current evidence reported shows a lack of development into the clinical testing phase; however the research does show future promise in the variation of different CAD systems.

Introduction

Even the best human observers make errors in the interpretation and classification of radiographs in making a diagnosis; be it a fracture, pathology or precursor to disease. These errors may be due to tiredness, inexperience, environmental disturbances or a combination of these.[1] As such, computers and software can potentially facilitate reducing these errors.[1] One of these facilitators is computer-aided detection (CAD), a technology designed to reduce observational oversights by using pattern recognition in order to bring attention to suspicious abnormalities within the image. CAD is designed to increase the sensitivity and specificity of a medical test.[2] CAD software has shown to increase diagnostic accuracy in many medical fields and thus helps physicians/radiologists to interpret medical images.[2] So far CAD has been integrated into some of the most common medical imaging examinations, for example:

- Mammography; improving the detection of micro calcifications.[3–5]
- Chest computed tomography (CT) scans; identifying pulmonary nodules via their density and shape[6]
- CT colonography; identifying colorectal polyps[7]
- Magnetic resonance imaging (MRI): prostate cancer screening[8]
- CT cardiac scans investigating coronary artery stenosis[9]
- Nuclear medicine whole body scans; where CAD identifying bone metastases[10]
- CT spinal imaging: detecting sclerotic bone metastases, and vertebral fractures in the spine[11,12]

These CAD programs have been shown to improve diagnostic accuracy and sensitivity in these fields, and are a clinically proven technology.[2–7] However it must be acknowledged that data exist that suggest CAD systems do not statistically improve accuracy of diagnosis,[13] and that CAD systems increase recall rates and reading times.[14] Although the majority of research shows positive results for CAD systems, there seems to be a lack of research regarding CAD software being used in musculoskeletal (MSK) medical imaging, this is especially important where inexperienced readers do more poorly at interpreting images in an acute trauma setting.[15]

A systematic review was undertaken to investigate the use of CAD software within MSK projection radiographic imaging, compared to the reference standard of current practice or a radiologists report. In addition, the review aimed to highlight possible evidence for further research.

* Corresponding author.
E-mail addresses: mg361@exeter.ac.uk (M. Gundry), K.M.Knapp@exeter.ac.uk (K. Knapp), R.M.Meertens@exeter.ac.uk (R. Meertens), J.R.Meakin@exeter.ac.uk (J.R. Meakin).

https://doi.org/10.1016/j.radi.2017.11.002
1078-8174/

Methods

This systematic review was carried out according to the guidance provided by the Cochrane Collaboration with regards to systematic reviews and diagnostic test accuracy studies,[16,17] whilst also utilising the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement.[18]

Eligibility criteria

The participants included any patients with suspected MSK abnormalities/pathologies/injuries from any age, gender, or background. The index test/intervention was any form of CAD applied to projection radiography including dual energy x-ray absorptiometry (DXA) and orthopantomography (OPT). Exclusions included any other medical imaging modalities such as: CT, nuclear medicine, MRI, ultrasound, mammography and colon imaging. Papers which discussed the technical aspect of CAD but did not test the software were also excluded. Additionally the use of CAD had to be the primary focus of the paper to be included. The reference standard/comparator was the current practice being utilised, in most cases this was the diagnostic report created by a radiologist. The primary outcome measures were sensitivity and specificity scores, and area under a curve (AUC) differences involving CAD as defined by the study. Therefore studies which did not include these measures of diagnostic test accuracy were excluded. Secondary outcomes were differences in interpretation time and any issues or errors within the CAD system.

All relevant study designs were included with the exclusion of ideas, opinions, case studies and editorials. Only studies published after 2004 were included, due to the nature of CAD technology which is constantly evolving and improving. Thus anything prior to 2004 would be obsolete and simplified, furthermore its results would be outdated more error prone. Only articles published in English were included.

Information sources

To ensure all relevant research was identified, a wide selection of databases were searched: EMBASE, HMIC, MEDLINE (Ovid) (including Journals@Ovid full text, Your Journals@Ovid, Ovid MEDLINE corrections, Ovid MEDLINE Daily updates) Global Health, AMED, PubMed, ISI Web of Science, TRIP and Science Direct. In addition, references cited from included papers that were not retrieved utilising the search strategy but were deemed as relevant were included and subjected to the same study selection and extraction criteria.

Searching strategy

For each database a search strategy was performed, this included keyword terms, synonyms, and AND/OR qualifiers. These were grouped via their index test (e.g. “computer aided detection” OR “software aided diagnosis”) or their target condition (e.g. musculoskeletal “AND bone”). See Appendix A and Appendix B for examples of the database searches.

Study selection and data extraction

All results were extracted to EndNote (Endnote x7.0.1 Bld 7212 and Endnote x7.5 Bld 9325), and all duplicates were removed from the results pool and recorded in the PRISMA flow diagram (2.1.3 2009) as shown in Fig. 1. Two independent reviewers double screened the remaining studies using the title and abstract, against the eligibility criteria. Any disagreements were debated over by the two reviewers, with a third independent reviewer arbitrating. The included papers were then screened for full text inclusion against the eligibility criteria by the same two independent reviewers. These results were again compared, and any disagreements discussed, with the third reviewer having the final decision of their eligibility. Each excluded full text was accompanied by a justification as to its exclusion (e.g. text not retrievable, not MSK, no sensitivity or specificity data). Prior to data extraction the extraction form was trialled on two of the included papers and modifications made prior to full extraction. This extraction form included data such as: title, date of publication, pathology, images/patients used, how patients/images were recruited, CAD sensitivity and specificity scores, details of the reference standard, differences in interpretation times, key conclusions, and miscellaneous comments by the author or the reviewer. A truncated version of this information is seen in Tables 1–3.

Risk of bias in individual studies

A protocol was developed and tested and modifications made prior to any data extraction, this mainly included introducing the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool[19] instead of the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) tool,[20] due to QUADAS-2 being more specialised in diagnostic accuracy studies.

Prior to data extraction the QUADAS-2 signalling questions were modified to be more relevant to the CAD centred response, this tool was then employed on all eligible studies for assessing the quality and presence of bias in the included papers. It was piloted by two independent reviewers, and then applied by two independent reviewers.

Analysis

The main primary outcome measure in all studies was sensitivity and specificity, this included AUC scores. A meta-analysis was considered but due the wide range of pathologies and CAD type likely to be discovered this would lead to high heterogeneity and thus a narrative review was more appropriate.

Results

Study selection

6253 studies were identified, 12 papers (0.19%) could not be accessed or retrieved despite several attempts and thus could not be included. Following the PRISMA flow diagram (Fig. 1) primary screening resulted in 149 papers remaining for full text review, 19 of the 149 (7.84%) were meditated on by a third researcher (JM) due to disagreement between the two reviewers. A total of 24 papers were included in the final data extraction, upon extraction two of these papers were excluded, one due to being a form of literature review[21] and cited the papers mentioned within the review so did not provide any additional information, and the second[22] on closer inspection was a technical modelling paper, both papers passed the inclusion criteria but upon investigation failed to provide any new or relevant information so were excluded from the final data extraction.

Study characteristics

A condensed version of the characteristics of the included studies is divided into their different CAD pathologies is shown in Tables 1–3. Of the final 22 studies; five utilised CAD to determine vertebral fractures; of which three investigated CAD in lateral chest

Please cite this article in press as: Gundry M, et al., Computer-aided detection in musculoskeletal projection radiography: A systematic review, Radiography (2017), https://doi.org/10.1016/j.radi.2017.11.002