



UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Advances in digital chest radiography: impact on reader performance

De Boo, D.W.

Publication date
2012

[Link to publication](#)

Citation for published version (APA):

De Boo, D. W. (2012). *Advances in digital chest radiography: impact on reader performance*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

Samenvatting



Samenvatting

Sinds de introductie van digitale thorax radiografie zijn er voortdurend verbeteringen geïntroduceerd op het gebied van dosis efficiëntie van de detectoren. Tevens maakte de digitale thorax radiografie de weg vrij voor de ontwikkeling van diverse processing tools. Dit proefschrift behandelt de effecten van de verschillende voordelen die digitale thorax radiografie biedt en dan vooral de invloed hiervan op de prestaties van radiologen.

In **hoofdstuk 2** wordt mobiele direct radiografie (DR) en mobiele computed radiografie (CR) vergeleken voor thorax foto's gemaakt aan bed bij patiënten opgenomen op een volwassen intensive care afdeling. De algehele beeldkwaliteit, de afgrensbaarheid van anatomische structuren en van lijnen / tubes werden beter beoordeeld met DR in vergelijking met CR. Zelfs DR met 50% dosis reductie ($DR_{50\%}$) was beter dan CR zowel met betrekking tot de afgrensbaarheid van mediastinale structuren als die van de lijnen / tubes. De beeldkwaliteit werd tijdens de separate beoordeling als gelijk beoordeeld. Echter tijdens de directe vergelijking werd de beeldkwaliteit van $DR_{50\%}$ in 87% van de gevallen als beter beschouwd dan die van CR. Interobserver variabiliteit voor de beoordeling van pathologie werd als surrogaat gebruikt om te onderzoeken of verbeterde beeldkwaliteit een effect zou hebben op de diagnostische prestaties. Alleen DR had een variabiliteit van 0.48, welke gezien wordt als klinisch acceptabel. $DR_{50\%}$ en CR hadden een lagere, vergelijkbare variabiliteit (respectievelijk, 0.39 en 0.33).

Hoofdstuk 3 richt zich op de effecten van "grey-scale reversal" op de detectie van kleine long noduli in de thorax radiografie. "Grey-scale reversal" is een simpel hulpmiddel dat met één muisklik beschikbaar is op elk PACS werkstation. Men weet vanuit optische fysiologie dat contrast perceptie beter is wanneer een donker object op een witte achtergrond wordt gepresenteerd. Wij vermoedden dat een geïnverteerde thoraxfoto ("bones black") de beoordelaar zou helpen in het detecteren van kleine noduli. Drie radiologen in opleiding en drie ervaren radiologen hadden echter geen baat bij "grey-scale reversal" voor de detectie van de noduli (gemiddelde diameter 13mm; mediaan 11mm). Wij hebben hieruit geconcludeerd dat "grey-scale reversal" geen bruikbaar hulpmiddel is voor de detectie van kleine long noduli.

Hoofdstuk 4 biedt een overzicht van de publicaties over computer-aided detection (CAD) voor de detectie van long noduli en T1 longcarcinomen ten tijde van de uitvoering van onze twee

observer studies. Zeer uitlopende resultaten worden gerapporteerd voor de verschillende prototypes en FDA goedgekeurde CAD systemen. De prevalentie van noduli was in alle studies hoger dan normaal. In vele studies werd slechts de stand-alone performance beoordeeld. Deze wordt echter sterk beïnvloed door prevalentie, detecteerbaarheid en grootte van de studie laesies. Net als bij de beoordeling van de CAD stand-alone performance, bepalen waarnemers en laesie selectie grotendeels de uitkomst bij observer studies. De ervaring van de waarnemers en distributie van detecteerbaarheid van de laesies zal het potentiële effect van CAD sterk beïnvloeden. Daarnaast presteert het ene CAD systeem anders dan het andere. Al deze aspecten maken het vergelijken van de tot nu gepubliceerde studies lastig en men dient voorzichtig te zijn met het trekken van algemene conclusies over CAD. Experts pleiten derhalve ook voor een uniforme vergelijking door middel van publiekelijk beschikbare databases met gevalideerde thorax foto's.

In **hoofdstuk 5** en **hoofdstuk 6** worden de twee huidig beschikbare FDA goedgekeurde CAD systemen onderzocht. Naast de daadwerkelijke analyse verschillen de systemen onderling zowel met betrekking tot de weergave van de CAD markeringen als de integratie in de dagelijkse praktijk.

Hoofdstuk 5 beschrijft de resultaten van de eerste observer studie waarin de invloed van CAD (Onguard 5.0; Riverain, Miamisburg, Ohio) op de detectie van T1 tumoren van deelnemers aan de Nederlands-Belgische longkanker screening studie (NELSON) wordt onderzocht. De sensitiviteit van CAD en die van ervaren radiologen was gelijk, echter het gemiddeld aantal fout positieve markeringen per foto was tien maal hoger voor CAD in vergelijking met die van de radiologen (2.4 vs. 0.24). Ondanks dat CAD tumoren markeerde die de radiologen initieel hadden gemist, leidde het gebruik van CAD niet tot een betere detectie van T1 tumoren. Van de 55 terecht positieve CAD markeringen werden er 5 tot 16 door de radiologen verworpen. Vooral terecht positieve CAD markeringen voor subtiele tumoren werden verworpen. Dit benadrukt het potentieel van CAD, maar tegelijkertijd ook dat gebruikers ervan moeite hebben om terecht positieve van fout positieve CAD markeringen te onderscheiden.

In **hoofdstuk 6** werd een studiegroep met andere studie laesies geselecteerd om het tweede FDA goedgekeurde CAD systeem (IQQA-Chest; EDDA Technology, Princeton Junction, NJ) te onderzoeken. De studiegroep bestond uit oudere patiënten waarvan de meerderheid rookte of gerookt had. Zowel ouderdom als roken leidt tot een toename van interstitiële longafwijkingen die in de literatuur

beschreven wordt als "dirty lungs". Kleine noduli met een lage detecteerbaarheid werden geselecteerd om de detectie capaciteit te testen. De sensitiviteit van onervaren waarnemers steeg significant van 39% naar 45% met gebruik van CAD als second reader ($p < 0.05$). Een niet significante stijging van het gemiddeld aantal fout positieve markeringen per foto (0.27 vs. 0.34) verhinderde echter een significante stijging van de figure of merit (0.71 vs. 0.71). De ervaren waarnemers presteerden beter dan de onervaren waarnemers, maar de verschillen zonder en met gebruik van CAD als second reader waren niet significant voor de sensitiviteit (50% vs. 51%), het gemiddeld aantal fout positieve per foto (0.16 vs. 0.21) en de figure of merit (0.84 vs. 0.87). Er werd 33% van de terecht positieve CAD markeringen verworpen en 40% van de fout positieve markeringen door de radiologen werd uitgelokt door een fout positieve CAD markering op dezelfde plaats. Dit liet wederom de moeite zien die gebruikers van CAD hebben om de terecht positieve van fout positieve CAD markeringen te onderscheiden.

Op basis van deze resultaten veronderstelden wij dat een gebrek aan vertrouwen in CAD een beter gebruik van de CAD resultaten beperkt. In **hoofdstuk 7** wordt onderzocht of training in het gebruik van CAD tot een betere detectie van long noduli zou leiden. De sensitiviteit van CAD was 59% en dit was iets lager in vergelijking met de sensitiviteit van 65% van de radiologen. Elke waarnemer kreeg individuele feedback na het beoordelen van een subset van foto's zonder en met CAD als second reader. De hypothese was dat deze feedback zou leiden tot een leercurve waardoor uiteindelijk de prestaties van de radiologen met CAD zouden verbeteren. De feedback resulteerde in een betere detectie van noduli zonder dat de CAD resultaten beter geïnterpreteerd werden. Dit gold voor beide manieren waarop CAD gebruikt kan worden: 1) met de mogelijkheid om markeringen te verwijderen die tijdens het beoordelen zonder CAD zijn gemaakt en 2) met alleen de mogelijkheid om extra markeringen te plaatsen tijdens het gebruik van CAD (add-on). In totaal werden door alle radiologen 32 terecht positieve CAD markeringen verworpen zonder verschil in resultaat tussen de eerste twee en laatste twee beoordelingssessies (16 vs. 16). Van de verworpen terecht positieve CAD markeringen was 78% voor slecht en zeer slecht detecteerbare noduli. Deze resultaten laten zien dat deze manier van training de moeite die gebruikers van CAD hebben om de terecht positieve van fout positieve CAD markeringen te onderscheiden niet verbetert.