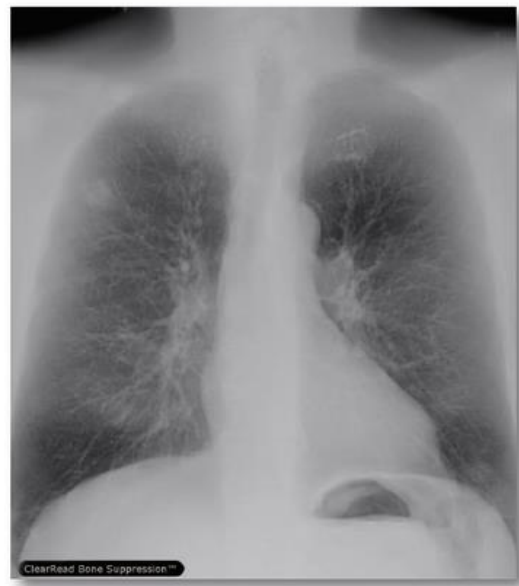


Fontys Paramedische Hogeschool

Opleiding Medisch Beeldvormende en Radiotherapeutische Technieken

De invloed van virtuele botsuppressie bij de evaluatie van thorax foto's

Is er een toegevoegde waarde bij de screening op maligniteiten?



Bron: <http://www.toyo.co.jp/>

Naam: M. R. H. (Marielle) Wouters
Studentnummer: 2143313

Onderzoeksverslag
Begeleider: Rob Vogels
7 Juni 2016

Voorwoord

In het kader van mijn afstudeerproject aan de opleiding Medisch Beeldvormende en Radiotherapeutische Technieken bij de Fontys Paramedische Hogeschool (FPH) heb ik dit onderzoeksverslag geschreven. Het onderzoek is uitgevoerd op de afdeling radiologie van het Maastricht Universiteit Medisch Centrum+ alwaar ik gedurende mijn duale opleiding werkzaam ben geweest.

Voor zijn eindeloze hulp wil ik Dr. U. Lalji bedanken. Hij was, ondanks zijn eigen promoveren, tijdens dit onderzoek mijn grote hulp en nam altijd de tijd mij te ondersteunen waar nodig. Daarnaast wil ik Drs. R. Korver – Steeman en B. Hendriks bedanken voor hun hulp bij het verzamelen van mijn data. Als praktijkopleiders wil ik René Pluymakers en Mariëlle Visser bedanken voor hun steun en begeleiding bij het proces. M. Bauwens en I. Houben wil ik bedanken voor hun hulp bij de statistiek. Als laatste maar zeker niet het minste wil ik Rob Vogels bedanken voor zijn hulp bij het schrijven van mijn scriptie en het contact op FPH.

Mariëlle Wouters, Juni 2016

Inhoudsopgave

Voorwoord	2
Samenvatting	4
Summary	5
Inleiding.....	6
Methode.....	8
Resultaten	11
Discussie	14
Conclusie	17
Literatuurlijst.....	18
Bijlage 1: Protocollen X-thorax.....	I
Bijlage 2: Scoreformulier	III
Bijlage 3: Berekening resultaten	IV

Samenvatting

Inleiding: Met behulp van een X-thorax kunnen afwijkingen in de thorax worden ontdekt. Echter, door superpositie van onder andere botstructuren kunnen potentiële afwijkingen gemist of verkeerd geïnterpreteerd worden. Aanvullend kan een CT-thorax worden geadviseerd. Omdat een CT-thorax hoger in dosis en kosten is, is een software ontwikkeld welke botstructuren op een X-thorax softwarematig elimineert (BSI). Het doel van dit onderzoek is om te bepalen of het gebruik van BSI, als aanvulling op een X-thorax, het aantal terechte adviezen voor een CT-thorax verhoogt en of BSI de zekerheid van de reviewer op het detecteren van maligne afwijkingen in de thorax bij een eerste screening vergroot.

Methode: 57 casussen zijn beoordeeld door een coassistent, een AIOS en een radioloog. Door middel van het invullen van één vragenlijst voor de X-thorax en één voor de BSI worden de sensitiviteit, specificiteit, PPV en NPV voor het adviseren van een CT-thorax bepaald. Om de gevonden resultaten met elkaar te kunnen vergelijken zijn ROC-curves opgesteld en AUC's bepaald. Om te kunnen bepalen of de reviewers door het gebruik van BSI minder twijfelen over de aanwezigheid van maligniteiten, worden Likertschalen ingevuld.

Resultaten: De sensitiviteit en NPV stegen bij alle reviewers door het gebruik van BSI. Voor de coassistent steeg in tegenstelling tot de andere reviewers, ook de specificiteit en de PPV. De gemiddelde AUC bleef nagenoeg gelijk, terwijl deze steeg voor de coassistent. Met name de coassistent leek zekerder te worden over de aanwezigheid van afwijkingen bij het gebruik van BSI.

Conclusie: Gebleken is dat een BSI het aantal terechte aanvullende CT-thorax adviezen voor een weinig ervaren reviewer verhoogt. Daarnaast wordt ook de zekerheid over de detectie van potentieel maligne afwijkingen voor de weinig ervaren reviewer door gebruik van BSI vergroot. Echter, om de betrouwbaarheid van de resultaten te vergroten is herhaling van de studie, met een grotere patiëntenpopulatie en een meer gedegen scholing gewenst.

Summary

Introduction: Using a chest X-ray malignant lesions can be discovered. Due to superposition of bony structures malignant lesions can be overlooked or misinterpreted. When in doubt, radiologist can suggest an additional chest CT. However, a chest CT is higher in radiation dose and costs. Therefore, software is developed to eliminate bony structures (BSI). The aim of this study is to investigate whether the use of BSI as an addition to chest X-ray could improve the number of justly requested chest CT-scans and if the confidence of the reviewers in detecting malignant lesions can be improved.

Material and methods: 57 cases are reviewed by an intern, a resident and a radiologist. Questionnaire were completed for the chest X-ray and BSI to determine the sensitivity, specificity, PPV and NPV for the request of additional chest CT's. To compare results, ROC-curves were obtained and AUCs were calculated. To determine if the confidence of the reviewers in detecting lesions increased Likert scales were used.

Results: Sensitivity and NPV increased for all the reviewers. For the intern, in contrast to the other reviewers, the specificity and the PPV increased. The average AUC stayed the same but increased for the intern. The reviewers were more confident about the appearance of 12 malignant lesions and were less confident about 6.

Conclusion: BSI as an addition to chest X-ray has a positive effect on the number of justly requested chest CT-scans for the reviewers with little experience. Hereby the interns confidence in detecting malignant lesions improves. However, repeating this study is recommended after longer training with BSI and including a larger patient population.

Inleiding

In Nederland worden jaarlijks ongeveer 12.000 mensen gediagnosticeerd met longkanker (1). Het is na prostaat-, huid- en darmkanker de meest voorkomende vorm van kanker (1). De persoonlijke, maatschappelijke en economische impact hiervan is groot (2). De symptomen die een patiënt kan ervaren variëren tussen afwijkingen in het hoestpatroon, versterkte sputumproductie, vermagering, haemoptoë en dyspnoe (1). Voor een goede prognose van longtumoren is het van belang dat elke, mogelijk maligne, afwijking in de longen in een vroeg stadium wordt ontdekt (3-5). Wanneer de patiënt klachten ervaart, zal hij in eerste instantie naar een huisarts gaan (3). De huisarts zal lichamelijk onderzoek uitvoeren en kan als aanvulling een röntgenfoto van de longen aanvragen (X-thorax) (3).

De X-thorax is een van de meest aangevraagde onderzoeken op een afdeling radiologie (6). Op een X-thorax kunnen afwijkingen groter dan vijf millimeter ontdekt worden (7). De X-thorax is snel gemaakt, goedkoop, geeft weinig vals positieve uitslagen en de patiënt ontvangt hierbij weinig ioniserende straling (7). Een nadeel van een X-thorax is dat een driedimensionaal orgaan op tweedimensionale wijze wordt weergegeven (7). Hierdoor worden de botstructuren in de thorax, het mediastinum, het hart en de longen in één beeld over elkaar heen geprojecteerd, wat ook wel superpositie wordt genoemd (7). Door superpositie van botstructuren kan het voorkomen dat afwijkingen gemaskeerd worden of dat radiologen niet zeker zijn over de interpretatie van de opname (5, 7-9). In geval van detectie van een verdachte afwijking of bij twijfel over de interpretatie van een opname zal een Computertomografie (CT) van de thorax als aanvullende beeldvorming worden geadviseerd.

Bij een CT-thorax wordt superpositie door het acquireren van dwarsdoorsneden geëlimineerd (10). De diagnostische accuratesse van een CT-thorax is mede hierdoor hoger dan die van een X-thorax waardoor de CT-thorax voor het ontdekken van maligne afwijkingen in de longen een betere modaliteit lijkt dan de X-thorax (4, 11-16). Echter mede door de hogere stralenbelasting, hogere kosten en het toedienen van jodiumhoudend contrastmiddel met bijkomende kans op allergische reacties en contrast geïnduceerde nefropathie (CIN), geniet een CT-thorax voor alsnog niet de voorkeur (10, 17, 18).

In een poging om de diagnostische accuratesse van de X-thorax te verhogen is software ontwikkeld welke botstructuren elimineert; de zogenaamde bone suppression imaging (BSI) (14, 15, 19, 20). Door het gebruik van BSI verdwijnt superpositie door botstructuren (14, 15, 19, 20). Originele X-thorax opnames worden door middel van software geanalyseerd en middels wiskundige algoritmes waarbij gebruik wordt gemaakt van de morfologie en densiteit van de structuren worden de projecties van botstructuren verwijderd (14, 15, 19-21). Bij de conversie van de X-thorax naar een BSI worden diverse filters gebruikt om de signaal-ruis verhouding en de contrast-ruis verhouding te optimaliseren, hierdoor presenteren structuren zich op een BSI anders dan op een X-thorax (19). Er bestaan verschillende vormen van maligne afwijkingen in de longen, waaronder gecalcificeerde noduli. Deze noduli vertonen op een X-thorax een zelfde densiteit als botstructuren (19). Echter omdat de BSI software gebruik maakt van de absorptie coëfficiënt en de morfologie van de botstructuren blijven deze in de conversie behouden (19).

Het is binnen het Maastricht Universitair Medisch Centrum⁺ (MUMC⁺) onduidelijk hoeveel terechte en onterechte CT-thorax opnames ter verduidelijking van potentiële maligniteiten op een X-thorax aangevraagd worden. Uit eerdere studies is gebleken dat door het gebruik van BSI het aantal aanvullende CT-thorax adviezen verhoogt. Of hierbij het aantal terechte en onterechte adviezen veranderd is niet eerder onderzocht (14, 15). Het is daarbij ook onduidelijk in hoeverre een BSI invloed heeft op de zekerheid van de reviewers over of een laesie reëel is of toch op superpositie berust. Het doel van deze studie is dan ook om te bepalen of het gebruik van BSI als aanvulling op een X-thorax het aantal terechte CT-thorax adviezen verhoogt en of BSI de zekerheid van de reviewers, op het detecteren van maligne afwijkingen in de thorax bij een eerste screening vergroot.

Methode

Middels dit retrospectieve, kwantitatieve onderzoek is onderzocht of het gebruik van een BSI als aanvulling op een X-thorax het aantal terechte CT-thorax adviezen verhoogt en of BSI de zekerheid van de reviewers, op het detecteren van maligne afwijkingen in de thorax bij een eerste screening vergroot. Dit onderzoek is uitgevoerd op de afdeling radiologie van het MUMC⁺.

Deelnemers

Patiënten van 18 jaar en ouder welke door de huisarts zijn doorverwezen voor een X-thorax en waarbij binnen 3 maanden na het vervaardigen van de X-thorax een CT-thorax ter verduidelijking is vervaardigd zijn geïnccludeerd. Van deze X-thorax diende een BSI aanwezig te zijn. Patiënten waarbij op de X-thorax verwijderbare voorwerpen zoals kettingen, piercings et cetera zichtbaar waren, of opnames welke niet gemaakt waren volgens het geldende protocol binnen het MUMC⁺ zoals beschreven in bijlage 1, zijn geëxcludeerd.

De geïnccludeerde patiënten zijn verdeeld in twee groepen. Een groep patiënten met potentieel maligne afwijkingen in de longen, mediastinum of lymfatisch stelsel van minimaal 5 millimeter gezien op de CT-thorax en een groep zonder potentieel maligne afwijkingen. De CT-thorax is in deze studie gebruikt als referentietest.

Uitkomstmaten

Om het aantal terechte en onterechte adviezen CT-thorax adviezen te kunnen bepalen worden de sensitiviteit, specificiteit, positief voorspellende waarde (PPV) en de negatief voorspellende waarde (NPV) bepaald. De sensitiviteit geeft aan welk percentage van de patiënten met potentieel maligne afwijkingen terecht een CT-thorax geadviseerd krijgt. De specificiteit geeft aan welk percentage van de patiënten zonder potentieel maligne afwijkingen terecht geen CT-thorax geadviseerd krijgt. PPV geeft aan welk percentage van positieve adviezen van CT-thorax, patiënten betreft met potentieel maligne afwijkingen en de NPV geeft aan welk percentage van de negatieve CT-thorax adviezen, patiënten betreft zonder potentieel maligne afwijkingen (22). Tot slot wordt middels Likertschalen bepaald of een BSI invloed heeft op de zekerheid van de reviewers, bij het detecteren van maligne afwijkingen in de thorax (23).

Meetinstrumenten

De X-thorax is gemaakt met een Siemens Ysio bucky, Erlangen, Duitsland. De X-thorax is bewerkt tot BSI door Softview (versie 2.4) van Riverain Technologies, Miamisburg, Ohio, Verenigde Staten. De CT-thorax opnames zijn gemaakt met een Siemens SOMATOM Definition Flash, Erlangen, Duitsland. Philips Brilliance 64 slice, Best, Nederland of Siemens SOMATOM Sensation 16, Erlangen, Duitsland. Alle gemaakte opnames zijn opgeslagen in het Picture Archiving and Communication System (PACS). De beelden werden beoordeeld op diagnostische beeldschermen Eizo RX340, Eindhoven, Nederland.

Dataverzameling

Visuele beoordeling

Scoreformulier

Door per casus een scoreformulier voor de X-thorax en een voor de BSI in te laten vullen konden de sensitiviteit, specificiteit, PPV en NPV worden berekend, volgens de formules uit bijlage 3. Aangegeven is of een aanvullende CT-thorax geadviseerd zou moeten worden en hoeveel potentieel maligne afwijkingen aanwezig waren met een maximum van drie. Deze afwijkingen zijn in een schematische figuur ingetekend om te kunnen correleren naar de verslagen van de CT-thorax. Om de zekerheid van de reviewers te kunnen bepalen zijn vierpunts Likertschalen ingevuld welke lopen van '1. Heel onzeker' tot '4. Heel zeker'.

Reviewers

De scorelijsten zijn om de klinische setting na te bootsen, ingevuld door een coassistent in het laatste jaar van de opleiding, arts in opleiding tot specialist radiologie (AIOS) in het vierde jaar van de opleiding met autorisatie voor verslaglegging van de X-thorax en een radioloog met aandachtsgebied in thorax met zeven jaar ervaring. De reviewers zijn niet betrokken geweest bij het maken van de opnames, hadden geen verdere klinische informatie en/of eerdere beeldvorming van de geïnccludeerde opnames, maar zijn wel geïnformeerd over het feit dat de dataset opnames mét en zónder maligne afwijkingen bevat.

Om een leercurve tijdens het beoordelen van de opnames te versnellen hebben de reviewers in de twee weken voorafgaand aan de studie een scholing gekregen in het beoordelen van BSI opnames. Hierin zijn de theoretische werking van de software uitgelegd en is geoefend met een training set van vijftien BSI opnames welke na afloop van de scholing zijn geëvalueerd. De inclusiecriteria van de BSI opnames in de training set waren hetzelfde als de inclusiecriteria van het gehele onderzoek.

De reviewers hebben de beelden onafhankelijk van elkaar en gerandomiseerd bekeken, waarbij de omstandigheden in de bejkrumte gelijk zijn gebleven. De reviewers hadden eerst alleen beschikking over de X-thorax in Posterior-Anterior (PA) en laterale richting met bijbehorende vragenlijst en direct hierna kwam de beschikking over de BSI met bijbehorende vragenlijst.

Data analyse

Statistische analyse

De statistische analyse is uitgevoerd met behulp van IBM Statistical Package for the Social Sciences (SPSS) versie 23, Amsterdam, Nederland.

Bepaling van het aantal terechte en onterechte CT-thorax adviezen ter aanvulling op een X-thorax

De sensitiviteit voor het doorverwijzen naar CT-thorax is bepaald door de terecht positieven (TP) te delen door de som van de TP en fout negatieven (FN). De specificiteit is bepaald door de terecht

negatieven (TN) te delen door de som van de TN en de fout positieven (FP). De PPV is bepaald door de TP te delen door de som van de TP en de FP. De NPV is bepaald door de TN te delen door de som van de FN en de TN. Een advies voor een aanvullende CT-thorax was TP als op het scoreformulier een positief advies werd gegeven voor een afwijking die in het originele verslag van de CT-thorax staat beschreven als potentieel maligne. Een advies was FP als een positief advies werd gegeven voor een afwijking die niet was beschreven in het verslag. Een advies was TN als een negatief advies werd gegeven aan een niet voor maligniteit verdachte patiënt. Een advies was FN als een negatief advies werd gegeven aan een patiënt die wel potentieel verdachte afwijkingen heeft.

Vergelijking van de relatie tussen de sensitiviteit en specificiteit van de X-thorax met en zonder BSI

Om de relatie tussen de sensitiviteit en de specificiteit te kunnen onderzoeken, werden Receiver Operating Characteristic curves (ROC-curves) opgesteld waarin de sensitiviteit voor het doorverwijzen naar een CT-thorax is uitgezet tegen één min de specificiteit (aspecificiteit). Om de X-thorax te kunnen vergelijken met de BSI zijn Area Under the Curve (AUC) berekent, waarbij een stijging van de AUC gelijk stond aan een stijging van het aantal terechte adviezen voor een CT-thorax.

Bepaling van de zekerheid van de detectie van potentieel maligne afwijkingen

Door beschrijvende statistiek toe te passen is bepaald of een BSI invloed heeft op de zekerheid van de reviewers over het detecteren van potentieel maligne afwijkingen. Dit is bepaald aan de hand van het aantal stappen dat de reviewers zekerder of onzekerder werden van de detectie. Hierbij stond een score van 1 gelijk aan een verhoging of verlaging van de zekerheid van de reviewers met 1 stap, bijvoorbeeld van 1. Heel onzeker, naar 2. Onzeker. Een score van 2 staat gelijk aan twee stappen enzovoorts.

Ethische aspecten

Tijdens dit onderzoek is gebruik gemaakt van retrospectieve data welke aanwezig is binnen het MUMC⁺. Hierdoor heeft dit onderzoek geen invloed op de gezondheid en het welbevinden van de patiënt waardoor deze studie niet binnen de Wet Medisch-wetenschappelijk Onderzoek valt. Patiënten welke binnen het MUMC⁺ in behandeling zijn, geven automatisch toestemming tot het gebruik van hun gegevens, tenzij hier bezwaar tegen wordt gemaakt (24).

Resultaten

Aan de hand van de gestelde criteria zijn 57 patiënten geïncludeerd voor dit onderzoek. Deze groep heeft een gemiddelde leeftijd van 64 jaar met een standaard deviatie van 14 jaar. De totale patiëntenpopulatie is opgesplitst in twee groepen. Te weten een 'groep met afwijkingen' (N=28) welke op de CT-thorax in totaal 39 potentieel maligne afwijkingen vertonen en een 'groep zonder afwijkingen' (N=29).

Tabel 1: Gegevens van de patiëntenpopulatie

	Totaal	Groep met afwijkingen	Groep zonder afwijkingen
Aantal (N=)	57	28	29
Man (N=)	33	18	15
Vrouw (N=)	24	10	14
Leeftijd in jaren (gem(SD))	64(14)	65(9)	63(17)
Totaal aantal afwijkingen	39*	39*	0

N = aantal, SD= standaard deviatie, gem = gemiddelde, * patiënten kunnen meerdere afwijkingen hebben

Bepaling van het aantal terechte en onterechte CT-thorax adviezen ter aanvulling op een X-thorax

Uit tabel 2 blijkt dat de radioloog meer CT-thorax opnames heeft geadviseerd dan andere reviewers. De coassistent zat met het aantal positieve adviezen onder het gemiddelde maar met de negatieve adviezen boven het gemiddelde. De radioloog zat met het aantal positieve adviezen boven het gemiddelde en met het aantal negatieve adviezen onder het gemiddelde. De adviezen van de AIOS komen nagenoeg overeen met het gemiddelde. Door het gebruik van BSI hebben alle reviewers één patiënt extra een terecht positief advies gegeven. De radioloog heeft echter drie patiënten door BSI ten onrechte doorverwezen voor een CT-thorax. De coassistent heeft twee fout positieve adviezen veranderd in terecht negatieve adviezen.

Tabel 2: weergave van de accuratesse van de adviezen voor een aanvullende CT-thorax van de verschillende reviewers waarbij de X-thorax wordt vergeleken met de BSI

Reviewer	X-thorax				BSI			
	Coas	AIOS	Rad	Gem	Coas	AIOS	Rad	Gem
Terecht positief	16	18	23	19	17	19	24	20
Fout positief	8	10	17	11	6	10	20	12
Terecht negatief	22	20	14	19	24	20	11	18
Fout negatief	11	9	3	8	10	8	2	7

Coas = coassistent, AIOS = arts assistent in opleiding tot specialisme, Rad = radioloog, Gem = gemiddelde van de reviewers

Tabel 3: Weergave van de sensitiviteit, specificiteit, PPV en NPV voor het adviseren van een CT-thorax voor de verschillende reviewers waarbij wordt vergeleken tussen de X-thorax zonder BSI en de X-thorax met BSI

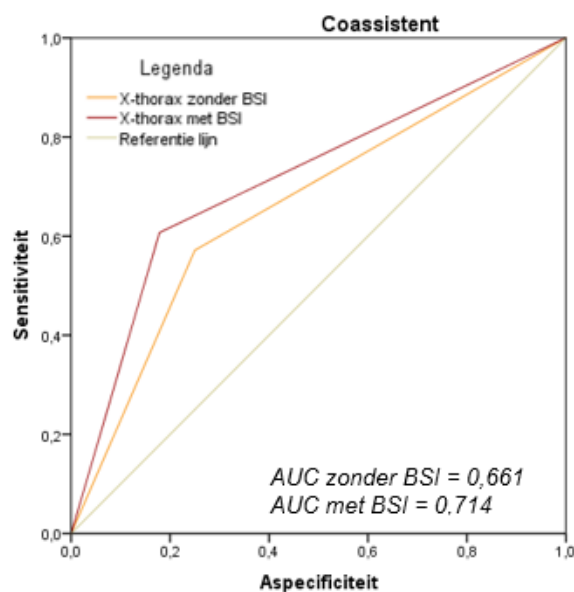
	X-thorax				BSI			
Reviewer	Coas	AIOS	Rad	Gem	Coas	AIOS	Rad	Gem
Sensitiviteit	59	68	88	73	63	71	92	75
Specificiteit	73	67	45	62	80	67	35	60
PPV	67	65	56	63	73	67	55	65
NPV	67	69	82	73	71	71	85	76

Coas = coassistent, AIOS = arts assistent in opleiding tot specialisme, Rad = radioloog, Gem = gemiddelde van de reviewers, PPV = positief voorspellende waarde, NPV = negatief voorspellende waarde

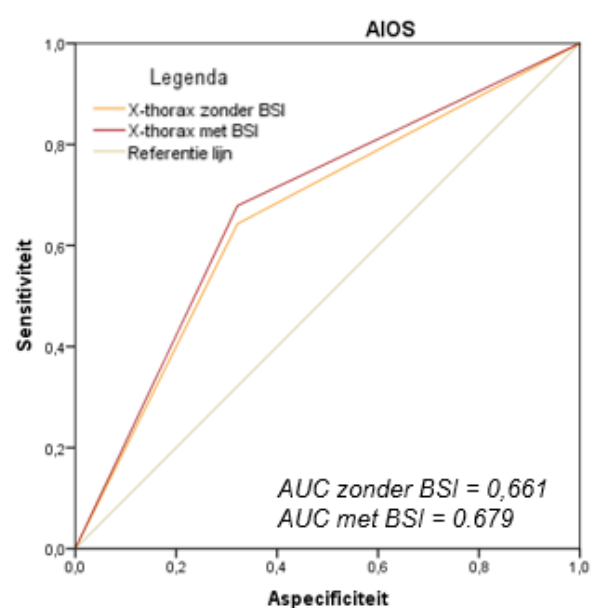
Uit tabel 3 blijkt dat de gemiddelde sensitiviteit, specificiteit, PPV en NPV nagenoeg gelijk blijven. De specificiteit is hoger dan het gemiddelde voor de coassistent. Daarbij steeg de specificiteit van de coassistent door het gebruik van BSI met 7% terwijl de specificiteit 10% daalde voor de radioloog. De sensitiviteit van de radioloog bleek 15% hoger dan het gemiddelde terwijl zijn PPV 10% lager was dan het gemiddelde.

Vergelijking van de relatie tussen de sensitiviteit en specificiteit van de X-thorax met en zonder BSI

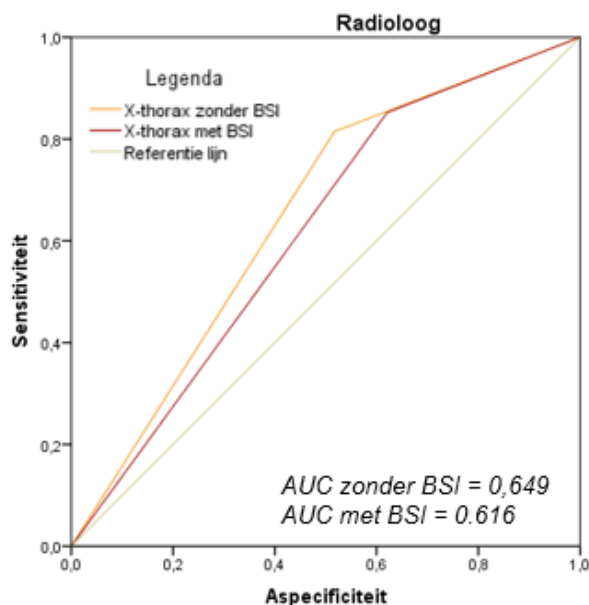
Figuur 1-3 laten de ROC-curves van de verschillende reviewers zien met bijbehorende AUC's. In figuur 4 is de ROC-curve met de gemiddeldes van alle reviewers te zien. De AUC van de coassistent steeg bij het zien van de BSI terwijl de AUC van de radioloog daalde. De AUC van de AIOS en de gemiddelde AUC bleven nagenoeg gelijk.



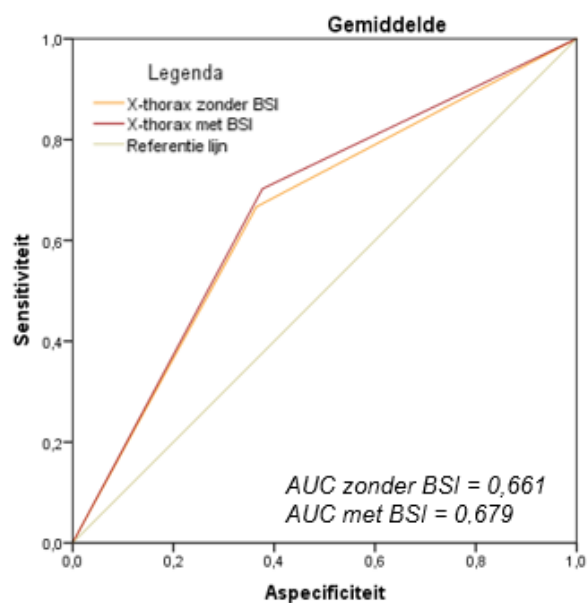
Figuur 1: ROC curve met bijbehorende AUC van de coassistent waarbij de sensitiviteit wordt uitgezet tegen de aspecificiteit



Figuur 2: ROC curve met bijbehorende AUC van de AIOS waarbij de sensitiviteit wordt uitgezet tegen de aspecificiteit



Figuur 3: ROC curve met bijbehorende AUC van de radioloog waarbij de sensitiviteit wordt uitgezet tegen de aspecificiteit



Figuur 4: ROC curve met bijbehorende AUC van het gemiddelde van de reviewers waarbij de sensitiviteit wordt uitgezet tegen de aspecificiteit

Bepaling van de zekerheid over de detectie van potentieel maligne afwijkingen

In tabel 3 is te zien dat de reviewers gemiddeld 12 keer zekerder werden over de aanwezigheid van een afwijking en 6 keer minder zeker. De coassistent en de radioloog zijn beide van 16 afwijkingen zekerder geworden terwijl de AIOS van maar vier afwijkingen zekerder is geworden. Daarbij zijn de AIOS en de radioloog beide van vier afwijkingen onzekerder geworden in tegenstelling tot de coassistent die 11 keer onzekerder is geworden. Als de reviewers veranderen van zekerheid gaat dit meestal met 1 stap.

Tabel 3: Verandering in zekerheid over de detectie van potentieel maligne afwijkingen

	Coas		AIOS		Rad		Gem	
Δ Zekerheid	++	--	++	--	++	--	++	--
1 Stap	10	8	2	3	15	4	9	5
2 Stappen	5	2	2	1	1	0	3	1
3 Stappen	1	1	0	0	0	0	0	0
Totaal	16	11	4	4	16	4	12	6

Coas = coassistent, AIOS = arts assistent in opleiding tot specialisme, Rad = radioloog, Gem = gemiddelde van de reviewers, ++ = reviewer wordt zekerder over de aanwezigheid van potentiële maligniteiten, -- = reviewer wordt minder zeker over de aanwezigheid van potentiële maligniteiten

Discussie

Middels dit onderzoek is geprobeerd uit te zoeken of het gebruik van een BSI als aanvulling op een X-thorax het aantal terechte CT-thorax adviezen verhoogt en of een BSI de zekerheid van de reviewers, op het detecteren van maligne afwijkingen in de thorax bij een eerste screening vergroot.

Dit is bepaald aan de hand van een patiëntengroep van 57 patiënten. Van deze patiënten vertoonde 28 patiënten in totaal 39 afwijkingen verdacht voor maligniteit op de CT-thorax welke als referentietest is gebruikt. Gebleken is dat de gemiddelde sensitiviteit, specificiteit, PPV en NPV voor het adviseren van een aanvullende CT-thorax na het zien van de BSI nagenoeg gelijk blijven.

Opvallend is dat de sensitiviteit van de radioloog 15% hoger was dan het gemiddelde wat maakt dat hij patiënten met potentieel maligne afwijkingen vaker dan de andere reviewers terecht een positief CT-thorax advies zal geven. Daarbij is het ook opvallend dat de specificiteit van de coassistent 15% hoger was dan het gemiddelde wat maakt dat ze patiënten zonder potentieel maligne afwijkingen vaker dan de andere reviewers terecht een negatief CT-thorax advies zal geven. Daarnaast steeg de specificiteit van de coassistent door het gebruik van BSI met 7% terwijl deze voor de radioloog daalde met 10%. Hieruit volgt dat een BSI voor de coassistent een positief effect heeft op het geven van een terecht negatief advies voor de CT-thorax. Opvallend is ook dat de PPV van de coassistent hoger was dan die van de radioloog wat inhoudt dat het advies dat de coassistent geeft, relatief vaker terecht was dan dat van de radioloog. Daarnaast steeg de PPV voor de coassistent door het gebruik van de BSI wat inhoudt dat het aantal terechte CT-thorax adviezen steeg terwijl de PPV voor de AIOS en de radioloog nagenoeg gelijk bleven. De NPV van de radioloog was daarentegen wel hoger dan het gemiddelde, wat inhoudt dat de negatieve adviezen die de radioloog geeft, vaker terecht zijn. Uit de AUC's is gebleken dat de relatie tussen de sensitiviteit en de specificiteit voor de coassistent verbeterde door een BSI. Dit houdt in dat de accuratesse voor het adviseren van een aanvullende CT-thorax voor de coassistent beter wordt door het gebruik van BSI. Bij de radioloog daarentegen daalde de AUC door het gebruik van BSI, hiermee wordt de accuratesse op het adviseren van een aanvullende CT-thorax voor de radioloog minder als hij gebruik maakt van een BSI. De gemiddelde AUC en die van de AIOS bleven nagenoeg gelijk. Gebleken is ook dat de reviewers over de aanwezigheid van gemiddeld twaalf afwijkingen zekerder zijn geworden en van zes afwijkingen onzekerder. Echter als een reviewer met 1 stap in zekerheid veranderd, is dit geen significant verschil. Een reviewer kan dan immers nog steeds 'onzeker' zijn over de aanwezigheid van de afwijking. Gebleken is wel dat de reviewers over de aanwezigheid van gemiddeld drie afwijkingen 2 stappen zekerder zijn geworden en dus van 'onzeker' veranderde naar 'zeker'. Dit terwijl ze over de aanwezigheid van maar één afwijking 2 stappen onzekerder zijn geworden. Hiervan wordt de coassistent over de aanwezigheid van vijf afwijkingen 2 stappen zekerder. Het lijkt dus dat vooral de coassistent zekerder wordt over de detectie van maligne afwijkingen door het gebruik van BSI. Echter is het verschil in zekerheid van de reviewers door het gebruik van BSI niet overtuigend te noemen.

Freedman et. al. (14) en Schalekamp et. al. (15) hebben studies uitgevoerd die overeen komen met de huidige studie. Echter, omdat benigne afwijkingen in de voorgaande studies zijn geëxcludeerd en de patiënten maar één maligne afwijking mochten hebben, zijn de resultaten van die studies niet geheel met deze studie te vergelijken. Het aantal fout positieve adviezen zal door de geoptimaliseerde patiëntenpopulatie in deze studies lager liggen dan het aantal fout positieven in de huidige studie. Hierdoor kan te verklaren zijn waarom in die studies een hogere AUC werd gevonden. Doordat in de huidige studie alle voorkomende ziektebeelden zijn geïnccludeerd, zijn de resultaten van de huidige studie klinisch breder toepasbaar zijn dan de studies van Freedman et. al. (14) en Schalekamp et. al. (15). Ook is in de vorige studies er voor gekozen om op een schaal van 0 – 100 aan te geven hoe reëel de kans is dat een afwijking aanwezig is, waarbij is gesteld dat bij een score groter of gelijk aan 50 een CT-thorax geadviseerd moet worden. In de huidige studie is de mogelijkheid gegeven om een CT-thorax te adviseren, los van het feit of een potentieel maligne afwijking aanwezig is. Doordat in de huidige studie een CT-thorax adviseren een aparte mogelijkheid was, is het beeld dat over de adviezen is verkregen betrouwbaarder dan dat wat Freedman et. al. (14) en Schalekamp et. al. (15) hebben onderzocht.

Het is sterk dat in het huidige onderzoek meerdere ziektebeelden zijn geïnccludeerd. Omdat in de praktijk een breed scala aan ziektebeelden voorkomt maakt het dat de klinische toepasbaarheid van dit onderzoek beter is dan dat van de studies van Freedman et. al. (14) en Schalekamp et. al. (15) waarbij enkel bewezen maligne afwijkingen geïnccludeerd zijn.

Daarbij is het goed dat verschillende niveaus van ervaring de opnames hebben beoordeeld. Dit is conform de kliniek waar verslaglegging van de opnames ook gebeurt door verschillende niveaus van ervaring. Ook dit vergroot de klinische toepasbaarheid van het onderzoek.

Gekozen is om een vierpunts Likertschaal te verkiezen boven een vijfpuntsschaal waardoor de reviewers een keuze moeten maken tussen 'onzeker over de aanwezigheid van een afwijking' en 'zeker over de aanwezigheid van een afwijking'. Voor deze studie is dit een goede keuze gebleken omdat zo met een grotere betrouwbaarheid onderzocht is kunnen worden hoe zeker de reviewers zijn over de aanwezigheid van potentieel maligne afwijkingen. Als voor een vijfpuntsschaal was gekozen hadden de reviewers de mogelijkheid om een middelste weinigzeggende optie te kiezen waardoor de kracht van een Likertschaal wordt verminderd.

Het is goed dat de klinische gegevens en/of de oudere opnames zijn afgeschermd tijdens het beoordelen van de opnames. Door blinding van het onderzoek kan er een betrouwbaar verschil bepaald worden tussen de X-thorax en de BSI. Klinische gegevens en/of oudere opnames zouden voor een extra variabele in de resultaten zorgen waardoor niet met zekerheid gesteld zou kunnen worden of het gevonden verschil zou komen door het gebruik van BSI of door het gebruik van klinische gegevens dan wel oudere opnames.

Helaas heeft het blinderen van het onderzoek ook een nadeel, aangezien klinische gegevens uitermate belangrijk zijn voor het interpreteren van afwijkingen op een opname. Mogelijk kan bij de beschikbaarheid van klinische gegevens zoals koorts een afwijking eerder als pneumonie worden geïnterpreteerd dan als mogelijke maligniteit. Het aantal FP kan bij het gebruik van klinische gegevens en/of oudere opnames mogelijk afwijken.

Een andere beperkende factor van dit onderzoek is de grootte van de patiëntenpopulatie (N=57). Doordat bij een kleine patiëntenpopulatie één kleine verandering in de resultaten grote gevolgen kan hebben op de sensitiviteit, specificiteit, PPV en NPV maakt dat de gevonden resultaten van dit onderzoek mogelijk niet betrouwbaar zijn. Als meer patiënten geïnccludeerd worden zullen de resultaten betrouwbaarder zijn waarmee met grotere zekerheid uitspraken kunnen worden gedaan over de inzetbaarheid van BSI in de praktijk.

Het was voorafgaand aan de studie bekend dat een gedegen scholing nodig was om zo een leercurve tijdens het beoordelen van de opnames te elimineren. Echter, door een gebrek aan data is een scholing gegeven met enkel 15 patiënten. Uit het onderzoek van Schalekamp et. al. (15) is gebleken dat een scholing met 40 patiënten minimaal was en mogelijk niet voldoende om een leercurve tijdens het beoordelen van de opnames te voorkomen. De kans bestaat dat door de suboptimale scholing de reviewers tijdens het beoordelen van de opnames nog in een leercurve zaten wat maakt dat de resultaten mogelijk niet betrouwbaar zijn.

Binnen het MUMC+ is recent een nieuwe versie van de BSI software geïnstalleerd wat nieuwe perspectieven tot vervolgonderzoek biedt. De verschillende software versies zouden met elkaar vergeleken kunnen worden, waarbij het voor de betrouwbaarheid van de studie noodzakelijk is een grotere patiëntenpopulatie te includeren en de scholing voorafgaand aan de studie uit te breiden met meerdere casussen. Ook is het mogelijk het onderzoek te herhalen met de nieuwe versie van de software.

Uit de resultaten is gebleken dat een BSI als aanvulling op een X-thorax een positief effect heeft op het aantal terecht positieve adviezen van alle reviewers. Echter blijkt een reviewer met veel ervaring door het gebruik van BSI ook meer fout positieve adviezen te geven. Hieruit valt aan te bevelen dat een BSI als aanvulling op een X-thorax ingezet kan worden voor reviewers met weinig ervaring en voor de meer ervaren reviewers bij twijfel over een mogelijke afwijking. Door het gebruik van een BSI zal de zekerheid van vooral de minder ervaren reviewer toe nemen. Echter, de resultaten zijn door een beperkte patiëntenpopulatie en een suboptimale scholing niet geheel betrouwbaar.

Conclusie

Het doel van deze studie was om te bepalen of het gebruik van BSI als aanvulling op een X-thorax het aantal terechte CT-thorax adviezen verhoogt en of BSI de zekerheid van de reviewers, op het detecteren van maligne afwijkingen in de thorax bij een eerste screening vergroot.

Gebleken is dat BSI het aantal terechte aanvullende CT-thorax adviezen voor een weinig ervaren reviewer verhoogt. Daarnaast wordt ook de zekerheid over de detectie van potentieel maligne afwijkingen voor de weinig ervaren reviewer vergroot. Echter, om de betrouwbaarheid van de resultaten te vergroten is herhaling van de studie, met een grotere patiëntenpopulatie en een meer gedegen scholing gewenst.

Literatuurlijst

1. Gommer AM (RIVM) PMR. Hoe vaak komt longkanker voor en hoeveel mensen sterven eraan? Bilthoven: RIVM; 2014
2. Nederland LA. Longziekten feiten en cijfers 2013. Amersfoort: Microweb EDU; 2013.
3. Merck & Co. I. Longen en luchtwegen: Diagnose Whitehouse Station, NJ, USA: Merck Sharp & Dohme Corp.; 2003
4. Lee J-S, Wang J-W, Wu H-H, Yuan M-Z. A nonparametric-based rib suppression method for chest radiographs. Computers & Mathematics with Applications. 2012;64(5):1390-9.
5. Turkington PM, Kennan N, Greenstone MA. Misinterpretation of the chest x ray as a factor in the delayed diagnosis of lung cancer. Postgraduate medical journal. 2002;78(917):158-60.
6. Radiological Society of North America IR. X-ray (Radiography) - Chest: RSNA; 2015
7. Hensen J, Jaarsveld-Bolman K, Gemenen SG-v, Dol-Jansen J, Dam T. Radiologie: Techniek en onderzoek. Amsterdam: Reed Business; 2011.
8. Plas Avd. X-thorax: Startpunt Radiologie; 2014
9. Shah PK, Austin JH, White CS, Patel P, Haramati LB, Pearson GD, et al. Missed non-small cell lung cancer: radiographic findings of potentially resectable lesions evident only in retrospect. Radiology. 2003;226(1):235-41.
10. Hakkert M, Tempelman G, Dam T. Computertomografie Techniek, onderzoek en stralingshygiëne: Springer Media B.V. ; 2015.
11. Toyoda Y, Nakayama T, Kusunoki Y, Iso H, Suzuki T. Sensitivity and specificity of lung cancer screening using chest low-dose computed tomography. Br J Cancer. 2008;98(10):1602-7.
12. Kligerman S, Cai L, White CS. The Effect of Computer-aided Detection on Radiologist Performance in the Detection of Lung Cancers Previously Missed on a Chest Radiograph. Journal of Thoracic Imaging. 2013;28(4):244-52.
13. White CS, Flukinger T, Jeudy J, Chen JJ. Use of a Computer-aided Detection System to Detect Missed Lung Cancer at Chest Radiography. Radiology. 2009;252(1):273-81.
14. Freedman MT, Lo SC, Seibel JC, Bromley CM. Lung nodules: improved detection with software that suppresses the rib and clavicle on chest radiographs. Radiology. 2011;260(1):265-73.
15. Schalekamp S, van Ginneken B, Meiss L, Peters-Bax L, Quekel LG, Snoeren MM, et al. Bone suppressed images improve radiologists' detection performance for pulmonary nodules in chest radiographs. European journal of radiology. 2013;82(12):2399-405.
16. Li F, Engelmann R, Armato III SG, MacMahon H. Computer-Aided Nodule Detection System: Results in an Unselected Series of Consecutive Chest Radiographs. Academic Radiology. 2015;22(4):475-80.

17. America RSoN. Radiation Dose in X-Ray and CT Exams 2015
18. Sharma SK, Kini A. Effect of nonionic radiocontrast agents on the occurrence of contrast-induced nephropathy in patients with mild-moderate chronic renal insufficiency: pooled analysis of the randomized trials. Catheterization and cardiovascular interventions : official journal of the Society for Cardiac Angiography & Interventions. 2005;65(3):386-93.
19. Technologies R. FAQ 2014
21. Novak R, Novak N, Gilkeson R, Mansoori B, Aandal G. A Comparison of Computer-Aided Detection (CAD) Effectiveness in Pulmonary Nodule Identification Using Different Methods of Bone Suppression in Chest Radiographs. J Digit Imaging. 2013;26(4):651-6.
22. Bakker E, Buuren Hv. Onderzoek in de gezondheidszorg. Nederland: Noordhoff Uitgevers B.B.; 2009.
23. McLeod SA. Likert Scale 2008
24. UMC+ M. Wetenschappelijk onderzoek 2015

Bijlage 1: Protocollen X-thorax

Protocol X-thorax

Thorax PA

Positionering:

- De patiënt staat goed rechtop, met de borstkast tegen de detector.
- De scapulae laten wegdraaien, door de schouders te laten zakken en de ellebogen naar ventraal te draaien.

Centrering:

- De bovengrens van het lichtvizier valt op C7.
- De centraalstraal is gericht op de mediaanlijn ter hoogte van Th7/Th8 (halverwege het corpus sterni).

Stralenrichting:

- PA.

Criteria:

- De thorax wordt symmetrisch afgebeeld: de afstand van de processus spinosus tot de claviculae is beiderzijds gelijk.
- De clavicula projecteren zich tussen het tweede en derde ribbenpaar.
- Scherpe afbeelding van de sinus pleurae.
- Bij een goede inspiratie dienen in de longvelden minimaal 10 dorsale ribbenparen te zijn geprojecteerd.
- De zwarting van het mediastinum is zodanig, dat door het mediastinum de contouren van de wervelkolom en de trachea, inclusief de bifurcatie, nog net te zien zijn.
- De longvaattekening is zichtbaar.

OPMERKINGEN:

- Deze opname kan ook bij uitzondering AP zittend (bv. Sterke lordose) gemaakt worden, of AP liggend in het geval van immobiliteit.
- Alle thoraxfoto's met maximale inspiratie maken. Onvoldoende inspiratie kan op de foto met name in de onderkwabben van de longen een hyperdensiteit weergeven, die longpathologie suggereert.
- Bij vraagstelling pneumothorax alleen inspiratie maken. Indien er een sterke verdenking is op een rand-pneu en deze is niet zichtbaar op een inspiratie-opname, kan eventueel in overleg met de radioloog een expiratie gemaakt worden.
- Bij vraagstelling vrije lucht, de patiënt minstens 10 minuten rechtop laten zitten alvorens de foto te maken.

Thorax LAT

Positionering:

- De patiënt staat met de linkerkzijde tegen de detector.
- De patiënt pakt met beide handen de beugel van de wandbucky boven het hoofd vast, of de handen op het hoofd leggen. De ellebogen naar elkaar toebrengen.
- Het bovenlichaam wordt iets naar voren gekanteld. Hierdoor worden de dorsale en ventrale zijde van de sinus pleurae op één horizontale lijn geplaatst.

Centrering:

- De centraalstraal is gericht ter hoogte van Th7/Th8 in het midden van de thorax.

Stralenrichting:

- Latero-mediaal.

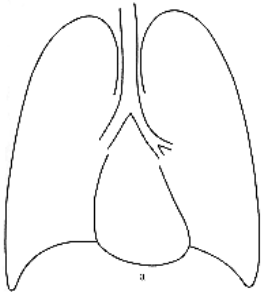
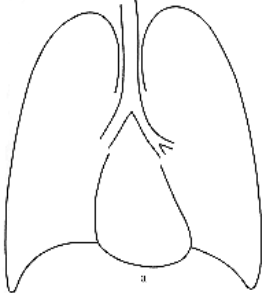
Criteria:

- De wervelkolom is lateraal afgebeeld.
- Scherpe afbeelding van de sinus pleurae.
- De longvaattekening is zichtbaar.
- Door de divergentie van de bundel worden de ribben rechts vergroot en iets dorsaal van de ribben links afgebeeld.

OPMERKINGEN:

- Deze opname kan ook bij uitzondering lateraal zittend (bv. Immobiliteit) gemaakt worden.
- Alle thoraxfoto's met maximale inspiratie maken. Onvoldoende inspiratie kan op de foto met name in de onderkwabben van de longen een hyperdensiteit weergeven, die longpathologie suggereert. Ook bij vraagstelling pneumothorax.

Bijlage 2: Scoreformulier

Reviewer: Coassistent, AIOS, Radioloog							
Casusnummer:							
X-thorax zonder BSI				X-thorax met BSI			
Geef hier met een nummer aan waar de afwijking zich bevindt.				Geef hier met een nummer aan waar de afwijking zich bevindt.			
 Figuur 5 Schematische tekening longen		1. Heel onzeker 2. Onzeker 3. Zeker 4. Heel zeker		 Figuur 1 Schematische tekening longen		1. Heel onzeker 2. Onzeker 3. Zeker 4. Heel zeker	
	Aanwezig?				Aanwezig?		
Afwijking 1	Ja / nee				Afwijking 1	Ja / nee	
Afwijking 2	Ja / nee				Afwijking 2	Ja / nee	
Afwijking 3	Ja / nee				Afwijking 3	Ja / nee	
Advies CT?	Ja/nee				Advies CT?	Ja/nee	

Bijlage 3: Berekening resultaten

Combinatiemogelijkheden resultaten

		Uitslag van de CT-thorax	
		Afwijking aanwezig	Afwijking afwezig
Uitslag X-thorax of BSI	Afwijking aanwezig	Waar positief (TP)	Vals positief (FP)
	Afwijking afwezig	Vals negatief (FN)	Waar negatief (TN)

Berekening sensitiviteit, specificiteit, PPV en NPV

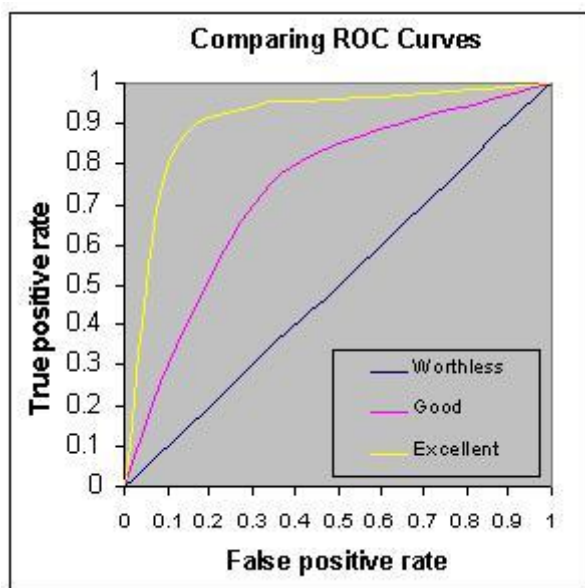
$$\text{Sensitiviteit} = \frac{TP}{(TP + FN)}$$

$$\text{Specificiteit} = \frac{TN}{(TN + FP)}$$

$$\text{Positief voorspellende waarde} = \text{PPV} = \frac{TP}{(TP + FP)}$$

$$\text{Negatief voorspellende waarde} = \text{NPV} = \frac{TN}{(TN + FN)}$$

ROC-curve



Addendum

Tijdens het uitvoeren van de studie is gebleken dat er een miscommunicatie tussen de opdrachtgever en de afstudeerder heerste. Derhalve is in onderling overleg met Dhr. Vogels, Dr. U. Lalji en Dhr. Pluijmakers afgeweken van het projectplan.

Volgens het projectplan zou onderzocht worden of door het gebruik van BSI de detectie van potentieel maligne afwijkingen accurater zou verlopen. Omdat het voor de kliniek van groter belang is om te weten of het aantal adviezen van een CT-thorax veranderd door het gebruik van een BSI is dit aangepast naar de vraag of het aantal terechte CT-thorax adviezen gaat verbeteren door het gebruik van BSI. Hierdoor is niet onderzocht wat de sensitiviteit, specificiteit, PPV en NPV is op het detecteren van maligne afwijkingen maar wat deze zijn op het adviseren van een CT-thorax.

Tijdens het berekenen of de zekerheid van de reviewers op het detecteren van maligne afwijkingen veranderd door het gebruik van BSI zijn we tot de conclusie gekomen dat een gemiddelde waarde over de Likertschaal niet de lading dekte. Daarom is er voor gekozen om middels beschrijvende statistiek te bepalen hoe vaak de reviewers zekerder worden en hoe vaak onzekerder over de detectie van potentieel maligne afwijkingen