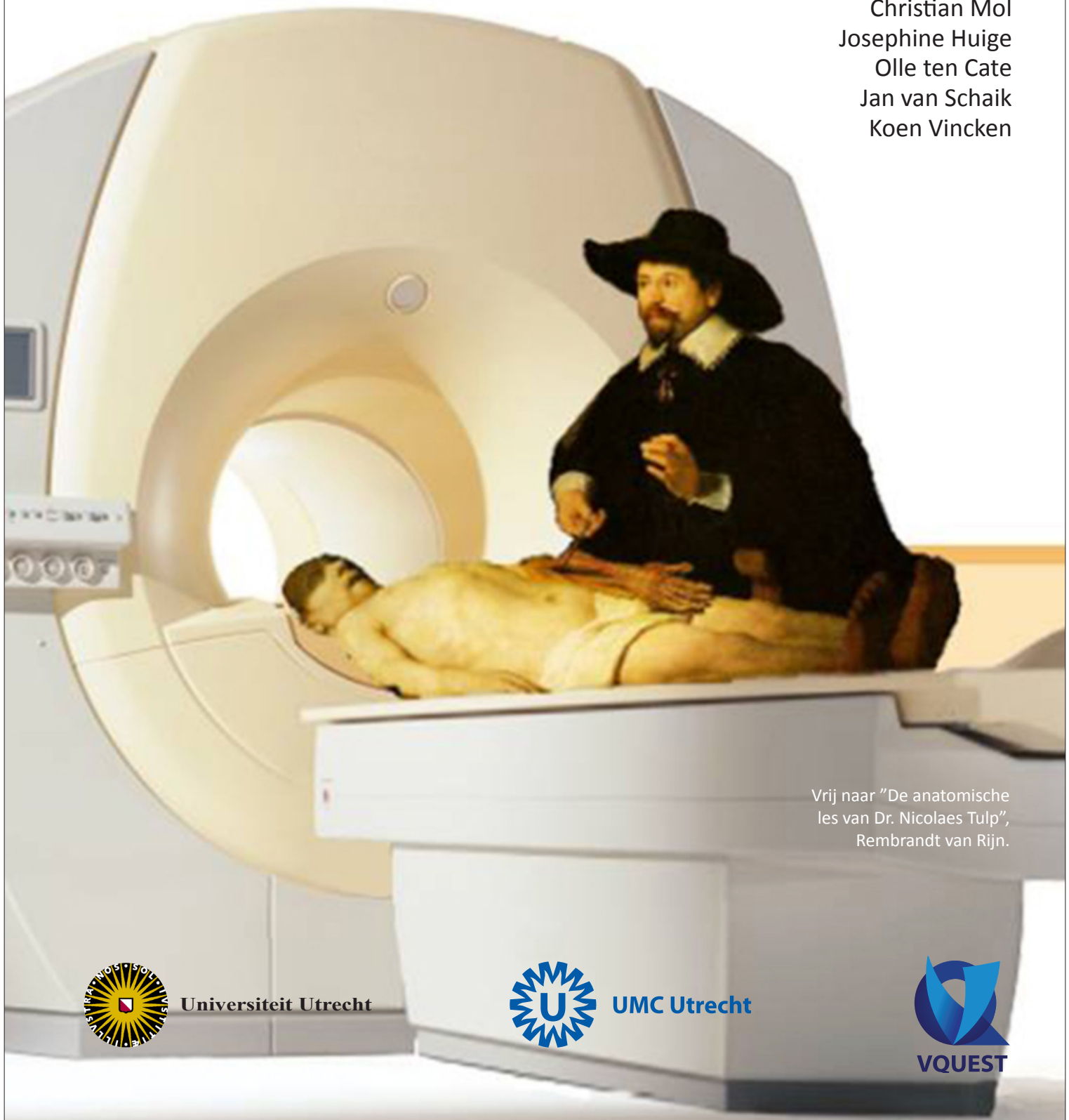


DIGITALE TOETSING VAN RADIOLOGISCHE BEELDINTERPRETATIE

Cécile Ravesloot
Anouk van der Gijp
Marieke van der Schaaf
Ronald Bleys
Christian Mol
Josephine Huige
Olle ten Cate
Jan van Schaik
Koen Vincken



Vrij naar "De anatomische
les van Dr. Nicolaes Tulp",
Rembrandt van Rijn.



Universiteit Utrecht



UMC Utrecht



VQUEST

Inhoud

1) Inleiding en begrippenlijst	5
2) Beeldinterpretatie in de radiologische praktijk	7
2.1. Vaardigheden om radiologische beelden te interpreteren	
2.2. Verschillen in denkprocessen tussen 2D en MPR beeldinterpretatie	
2.3. Verschillen in kijk- en zoekgedrag tussen 2D en MPR beeldinterpretatie	
2.4. Ontwikkeling van vaardigheden in radiologische beeldinterpretatie	
3) Optimalisering van de kwaliteit van toetsing van beeldinterpretatie	17
3.1. Inleiding	
3.2. Onderzoeksmethoden	
3.3. Resultaten	
3.4. Samenvatting	
4) Stappenplan voor het afnemen van een digitale toets voor vaardigheden in radiologische beeldinterpretatie	24
4.1. Inleiding	
4.2. Toetsvragen maken	
4.3. Toets afnemen	
4.4. Toets nakijken	
5) Conclusie	30
6) Referentielijst	31
Appendix Studenthandleiding VQuest	33

1. Inleiding

In het onderzoeksproject “Toetsing in Beeld” (2011-2014), gesubsidieerd door SURF, is onderzocht of en op welke wijze de kwaliteit van de toetsing van radiologische beeldinterpretatie met behulp van multiplanar reformatted (MPR) beelden verhoogd zou kunnen worden.

De radiologie is het vakgebied dat zich van origine bezig houdt met het vaststellen van ziektes (diagnosticeren) aan de hand van beelden van de inwendige mens, radiologische beelden. De vaardigheid die komt kijken bij het interpreteren van beelden en die leidt tot een conclusie (diagnose) wordt beeldinterpretatie genoemd. In het geval van de radiologie spreken we van radiologische beeldinterpretatie. Dankzij digitale en technische ontwikkelingen kunnen tegenwoordig in enkele seconden honderden aaneensluitende dwarsdoorsnedenbeelden van het menselijk lichaam worden gemaakt. Dit maakt het mogelijk om deze volume datasets te bekijken in stack mode, waarbij de arts als het ware door de patiënt heen scrollt. Deze beelden worden ook wel multiplanar reformatted (MPR) beelden genoemd. Ter vergelijking, zo’n 10 jaar geleden werden radiologische dwarsdoorsnedenbeelden bekeken als naast elkaar afgedrukte plaatjes van het menselijk lichaam (2D). Hoewel digitale MPR beelden volledig zijn geïntegreerd in de klinische praktijk, vindt de toetsing van de vaardigheid in beeldinterpretatie in de meeste geneeskunde- of medisch specialistische opleidingen nog plaats aan de hand van 2D beelden. Er is momenteel weinig bekend over de invloed van toetsing aan de hand van 2D en MPR beelden op de validiteit van toetsing van radiologische beeldinterpretatievaardigheden.

De hypothese onderliggend aan dit project is dat bij digitale toetsing met MPR beelden, vaardigheden om radiologische beelden te interpreteren meer valide en efficiënter kunnen worden beoordeeld dan bij digitale toetsing met 2D beelden. De evidentie hiervoor is onderzocht in een reeks onderzoeken waarbij data zijn verzameld tijdens diverse radiologietoetsen bij studenten van verschillende expertise niveaus (studenten geneeskunde en coassistenten in het UMC Utrecht). In dit project is gewerkt met het digitale toetsprogramma VQuest. VQuest is speciaal ontwikkeld voor het toetsen met zowel 2D als MPR beelden (zie appendix voor een handleiding). Zowel 2D en MPR beeldvragen werden onderzocht op verschillende aspecten van kwaliteit. Daarnaast werd inzicht verkregen in wat radiologische beeldinterpretatie behelst.

In hoofdstuk 2 wordt het begrip radiologische beeldinterpretatie gedefinieerd: wat houdt het in en hoe verloopt de ontwikkeling van deze vaardigheid? Ook wordt ingegaan op verschillen tussen 2D en MPR beeldinterpretatie. In hoofdstuk 3 worden de resultaten van het onderzoek naar de kwaliteit van de 2D en MPR beeldvragen besproken. Tot slot worden in hoofdstuk 4 de stappen voor het afnemen van een digitale toets naar vaardigheden in beeldinterpretatie besproken aan de hand van praktijkvoorbeelden. Hierbij worden tips en tricks voor het maken van beeldvragen geformuleerd.

Alle verrichte studies waarover wordt gerapporteerd zijn ethisch getoetst door de Nederlandse Vereniging voor Medisch Onderwijs. Van alle deelnemers is informed consent verkregen.

Begrippenlijst

2D beeld: Een 2D beeld is één geselecteerde doorsnede uit een MPR beeld. Hierbij is het niet mogelijk om het beeld in andere richtingen te bekijken.

2D en MPR beeldvraag: Dit is een toetsvraag aan de hand van een 2D- respectievelijk MPR beeld.

Aanwijsvraag: Toetsvraag waarbij de deelnemer gevraagd wordt een structuur of afwijking in een beeld te lokaliseren en te markeren.

Alternatievenlijst: Lange lijst met antwoorden voor longmenuvraag.

Beeldbank: Verzameling van beelden die geschikt zijn voor het maken van beeldvragen.

Beeldinterpretatie: Het beoordelen van beelden leidend tot een conclusie over de betekenis van het beeld.

Benoemvraag: Toetsvraag waarbij de deelnemer een in een beeld gemarkeerde structuur wordt gevraagd te identificeren.

Coupe: één doorsnede uit een MPR beeld.

Diagnose: Vastgesteld ziektebeeld.

Diagnosevraag: Toetsvraag waarbij de student gevraagd wordt aan de hand van een radiologisch beeld de meest waarschijnlijke diagnose te stellen.

Dicom-bestanden: Bestandsformat voor radiologische beelden.

Differentiaal diagnose: Lijst van mogelijke *diagnoses* in volgorde van waarschijnlijkheid (opgesteld op grond van bijvoorbeeld een bepaalde klachtenpresentatie of afwijkend radiologisch beeld).

Longmenuvraag: Toetsvraag waarbij de deelnemer uit een zeer lange lijst met antwoorden (*alternatievenlijst*) het juiste antwoord kan kiezen. Deze lijst bevat bij voorkeur meer dan 1000 opties en er kan middels woordherkenning in gezocht worden door een aantal letters van het gekozen antwoord in te typen.

Multiplanar reformatted (MPR) beeld: Dit beeld betreft een volumedataset, bestaande uit vele opeenvolgende dwarsdoorsneden van het menselijk lichaam. De beoordelaar kan door de dwarsdoorsneden heen scrollen in meerdere richtingen. Dit in tegenstelling tot een 2D beeld waarbij één doorsnede geselecteerd is uit de volumedataset.

Overlay: Gesegmenteerde structuur in het beeld door middel van het 'inkleuren' van de structuur in 2D of in een volumedataset. Hiermee kunnen bijvoorbeeld aanwijsvragen worden nagekeken.

Radiologisch beeld: Beeld dat met behulp van een radiologische techniek wordt gemaakt van (een deel van) het menselijk lichaam. Voorbeelden van radiologische beelden zijn: echogram (beeld vervaardigd met echografie), röntgenfoto (2D beeld vervaardigd met röntgenstraling), CT-scan (dwarsdoorsnedenbeeldvorming met behulp van röntgenstraling) en MRI-scan (dwarsdoorsnedenbeeldvorming met behulp van magnetische resonantie).

Radiologische beeldinterpretatie: Het proces van het interpreteren van radiologische beelden, leidend tot een conclusie, bijvoorbeeld in de vorm van een (differentiaal) diagnose.

Stack: serie van doorsnede beelden (die samen een MPR beeld vormen) waar men doorheen kan scrollen.

Summatieve toetsing: Toetsing waarvan het toetsresultaat consequenties heeft voor de studievoortgang van de studenten.

2. Beeldinterpretatie in de radiologische praktijk

2.1 Vaardigheden om radiologische beelden te interpreteren

2.1.1 Inleiding

Het interpreteren van radiologische beelden is fundamenteel voor de expertiseontwikkeling van radiologen. In het algemeen kunnen bij het interpreteren van beelden drie componenten worden onderscheiden; ‘perceptie’, ‘analyse’ en ‘interpretatie’ (Rogers 1996). De radioloog moet de radiologische afwijkingen verzamelen door ze te zien en te herkennen (perceptie). Tevens moeten deze bevindingen uiteengeerafeld worden om ze te begrijpen en te kunnen beoordelen: wat zijn de kenmerken van de afwijkingen of welke kenmerken ontbreken juist (analyse). Daarbij moeten deze bevindingen geïntegreerd worden en vertaald worden in een differentiaaldiagnose (interpretatie) en/of een advies voor behandeling of vervolgonderzoek. Een differentiaaldiagnose is een lijst met diagnoses op volgorde van waarschijnlijkheid in een specifiek geval.

De laatste jaren is de dagelijkse praktijk van beeldinterpretatie veranderd als gevolg van digitalisering en de introductie en ontwikkeling van MPR beelden. Omdat de beoordeling van MPR beelden om andere perceptuele en cognitieve activiteiten vraagt in vergelijking met de beoordeling van 2D beelden (Hegarty 2010), is er een verschuiving ontstaan in de benodigde radiologische expertise. Om de expertise van radiologen in de opleiding te garanderen is het van groot belang om de training en toetsing van radiologen in opleiding bij deze ontwikkelingen aan te laten sluiten. Dat geldt ook voor andere beroepen waarin beeldinterpretatie een belangrijke rol speelt. Om dit te kunnen bewerkstelligen is het belangrijk eerst vast te stellen welke kennis en vaardigheden nodig zijn om beelden te kunnen interpreteren.

Deze paragraaf rapporteert over een onderzoek naar kennis en vaardigheden die nodig zijn om radiologische beelden te kunnen interpreteren. Het doel van het onderzoek was om een raamwerk op te stellen dat de benodigde kennis en vaardigheden weergeeft. Voor verdere details van het onderzoek wordt verwezen naar Van der Gijp et al. (2014).

2.1.2. Onderzoeksmethode

Design

Voor het in kaart brengen van de benodigde kennis en vaardigheden om radiologische beelden te kunnen interpreteren zijn meerdere onderzoeksmethoden gebruikt. Het onderzoek bestond uit drie opeenvolgende fasen: een literatuurstudie, expertinterviews en een hardopdenkonderzoek.

Literatuurstudie

De literatuurstudie betrof een search in meerdere databases (PsychInfo, PubMed en Embase) en boeken. Kennis en vaardigheden die op basis van de literatuur belangrijk bleken voor beeldinterpretatie (Van der Gijp 2014) werden geëxtraheerd en opgenomen in de eerste versie van het raamwerk. De formulering en definities van de elementen werden na consensus in het onderzoeksteam opgesteld. De eerste versie van het raamwerk fungeerde als input voor de expertinterviews.

Expert interviews

Een selectie van experts werd individueel geïnterviewd over het raamwerk. Het expertteam bestond uit vier radiologen, een aios radiologie, twee onderwijskundigen, een cognitief psycholoog en een neuropsycholoog. De experts waren geselecteerd op basis van hun ervaring en expertise in hun domein en op basis van de affectie met radiologisch onderwijs. De radiologen hadden ervaring met het onderwijzen van beeldinterpretatie aan aios of medisch studenten. De niet-radiologen waren betrokken in onderzoek in radiologisch onderwijs. Deze selectie van experts maakte het mogelijk om de beeldinterpretatietaak vanuit verschillende invalshoeken te bediscussiëren. De interviews leidden tot aanpassingen in het eerste raamwerk. Het aangepaste raamwerk (de tweede versie) diende als codeerschema voor de analyse van het hardopdenkonderzoek.

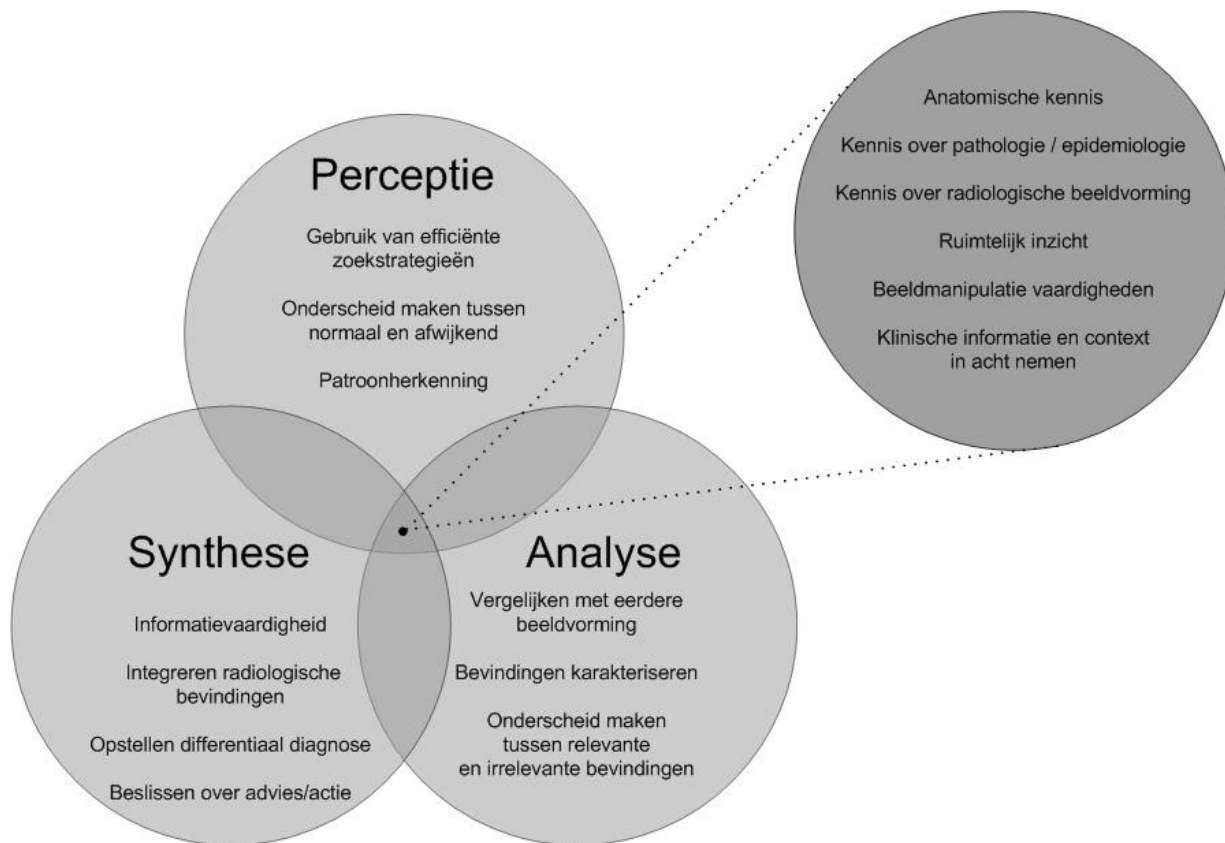
Hardopdenkonderzoek

In het hardopdenkonderzoek, werden medisch studenten gevraagd om hardop te denken tijdens het interpreteren van 2D en MPR beelden. Het betroffen tien vierde- tot zesdejaars geneeskundestudenten uit het UMC Utrecht, die allen het coschap radiologie hadden gevolgd en succesvol hadden afgesloten. De deelnemers kregen vier MPR casus en vier 2D casus. Elk van de vier domeinen in het coschap (neuroradiologie, thoraxradiologie, abdominale radiologie en musculoskeletale radiologie) was vertegenwoordigd in een 2D casus en in een MPR casus. De studenten kregen de opdracht om de beelden te beoordelen en hun gedachten hardop te verwoorden. De verwoordingen van de deelnemer werden opgenomen. Vervolgens werd de tekst uitgewerkt en gecodeerd met behulp van het codeerschema (de tweede versie van het raamwerk). De eerste vijf deelnemers werden door twee onderzoekers gescoord. Indien zij bemerkten dat een item uit het raamwerk niet goed aansloot bij de observaties uit de praktijk werd dit in het onderzoeksteam besproken en werden er zo nodig aanpassingen gedaan in het raamwerk. Dit leidde tot de uiteindelijke versie van het raamwerk.

2.1.3. Resultaten

In de literatuur op het terrein van radiologische beeldinterpretatie blijken twee stromingen te bestaan: een 'perceptuele' en een 'cognitieve' stroming (Beutel 2000), welke voornamelijk gebaseerd zijn op respectievelijk oogbeweging (eye-tracking) studies (Krupinski 1996; Kundel 1972, 1978, 1996; Nodine 1987, 2002; Cooper 2009) en studies die het hardopdenken tijdens het interpreteren van beelden analyseren (verbale protocollen) (Rogers 1996; Azevedo 1997; Lesgold 1988; Raufaste 1998; Morita 2008). Eye-tracking studies laten zien dat verschillen in beeldinterpretatievaardigheid berusten op perceptuele activiteiten (zoals zoekstrategieën en fixaties van de ogen). Hardopdenkonderzoeken rapporteren verschillen in cognitieve activiteiten, zoals redeneringen, evaluatie van radiologische bevindingen en het gebruik van verschillende soorten kennis (bijvoorbeeld anatomie en techniek). Dergelijke studies richtten zich tot nu toe op de beoordeling van conventionele röntgenfoto's (2D beelden); er zijn nog maar weinig studies verricht naar de vaardigheden die nodig zijn voor het beoordelen van MPR beelden.

De studie resulteerde in een raamwerk bestaande uit kennis en vaardigheden die een rol spelen tijdens de interpretatie van radiologische beelden. De drie componenten van beeldinterpretatie die in de literatuur beschreven waren (perceptie, analyse en interpretatie) werden ook in de praktijk van radiologische beeldinterpretatie onderscheiden door de experts in de interviews. Tevens waren ze herkenbaar in de hardopdenkstudies. Omdat in de derde component naast interpretatie ook andere vaardigheden belangrijk bleken, zoals het nemen van een beslissing of het geven van een advies is gekozen voor een andere, bredere term: synthese. Sommige kennis en vaardigheden bleken in meerdere componenten een rol te spelen. Om dit weer te geven zijn deze elementen in een vierde categorie in het raamwerk ondergebracht. Het resulterende raamwerk is weergegeven in figuur 1. Definities zijn samengevat in tabel 1.



Figuur 1: Raamwerk met elementen die nodig zijn voor digitale radiologische beeldinterpretatie. Vanuit het Engels vertaald uit: A. van der Gijp et al. (2014), 'Interpretation of radiological images: towards a framework of knowledge and skills'. *Advances in Health Sciences Education*.

Perceptie	
Gebruik van efficiënte zoekstrategieën	Het kunnen toepassen van efficiënte zoekstrategieën, zoals global zoeken, systematisch zoeken en hypothese-gestuurd zoeken. Zoekstrategieën kunnen achtereenvolgens en afwisselend toegepast worden tijdens het interpreteren van een radiologisch beeld.
Onderscheid maken tussen normaal en afwijkend	Het kunnen onderscheiden van normale bevindingen (en normaalvariatie) en abnormale bevindingen.
Patroonherkenning	Het direct en onbewust herkennen van de diagnose, gebaseerd op vergelijkbare patronen in het geheugen.

Analyse	
Vergelijken met eerdere beeldvorming	Radiologische bevindingen vergelijken met de bevindingen op eerdere beeldvormende onderzoeken van de patiënt.
Karakteriseren van bevindingen	Kenmerken (densiteit, vorm, contour) van bevindingen evalueren, indien nodig m.b.v. postprocessing.
Onderscheid maken tussen relevante en irrelevante bevindingen	Onderscheid maken tussen klinisch relevante bevindingen en bevindingen die niet klinisch relevant zijn.

Synthese	
Informatievaardigheid	Vaardigheden voor het doelmatig zoeken naar informatie in een system, waarin informatie is opgeslagen en weergegeven (bijvoorbeeld informatie in boeken of op internet).
Integreren van radiologische bevindingen	Radiologische bevindingen met elkaar in verband kunnen brengen (bijvoorbeeld een longnodule en een mediastinale lymfeklier)
Opstellen differentiaal diagnose	Het genereren van een lijst met mogelijke diagnoses in volgorde van waarschijnlijkheid. Dit omvat zowel het genereren als het verwerpen van diagnoses en het classificeren van ziektebeelden in een type of stadium.
Beslissen over advies/actie	Beslissingen nemen of het nodig is om een advies te geven (bijvoorbeeld aanvullende onderzoeken, vervolgonderzoeken) of actie te ondernemen (bijvoorbeeld de aanvrager opbellen)

Niet component-gerelateerd	
Anatomische kennis	Kennis van de anatomie, in het bijzonder de anatomische structuren zoals afgebeeld op radiologische beeldvorming.
Kennis over pathologie/epidemiologie	Kennis over pathologie (ziektebeelden) en in het bijzonder pathologie zoals afgebeeld op radiologische beeldvorming. Kennis van het klinisch beeld, behandelingen en het vóórkomen van ziektebeelden.
Kennis over radiologische beeldvorming	Kennis over de totstandkoming van de beelden en het effect van de radiologische beeldvormende technieken op de beelden.
Ruimtelijk inzicht	Het vermogen een 3D beeld te maken van een 2D weergave, dit 3D beeld te kunnen roteren en de relatie met andere weergaven van het beeld te kunnen analyseren.
Beeldmanipulatie vaardigheden (scrollen, wisselen van richting of weefselvenster)	Kennis hebben van het effect van beeldmanipulatie op de weergave van het beeld en in staat zijn te kiezen voor de optimale weergave voor het beoogde doel, bijvoorbeeld door het scrollen of aanpassen van weefselvenster of richting.
Klinische informatie en context in acht nemen	Het gebruik van klinische informatie (uit de aanvraag, het patiëntendossier en interactie met de patiënt) of klinische context (type ziekenhuis, aanvragend specialist) voor de interpretatie van het beeld.

Tabel 1: Definities van de elementen van beeldinterpretatie. Vanuit het Engels vertaald uit: A. van der Gijp et al. (2014), 'Interpretation of radiological images: towards a framework of knowledge and skills'. *Advances in Health Sciences Education*.

2.1.4 Praktische bruikbaarheid

Het raamwerk met kennis en vaardigheden voor 2D en MPR beeldinterpretatie, dat beschreven is in paragraaf 2.1, kan voor verschillende doeleinden worden toegepast in de onderwijspraktijk. Dat geldt zowel voor het ontwikkelen van een curriculum of cursus als voor het toetsen van kennis en vaardigheden in radiologische beeldinterpretatie. De elementen uit het raamwerk kunnen bijvoorbeeld gebruikt worden om leerdoelen op te baseren. Per ziektebeeld zou bijvoorbeeld een leerdoel kunnen worden geformuleerd, waar items van kunnen worden afgeleid. Zodoende kunnen een toetsmatrijs en toetsitembank voor beeldinterpretatievaardigheden worden opgebouwd. Ook zou het kunnen dienen als leidraad bij het inrichten van een les of lesmateriaal.

2.2. Verschillen in denkprocessen bij 2D en MPR beeldinterpretatie

2.2.1 Inleiding

In de voorgaande paragraaf is beschreven welke kennis en vaardigheden nodig zijn voor het interpreteren van 2D en MPR radiologische beelden. In deze paragraaf wordt ingegaan op het mogelijke verschil in benodigde kennis en vaardigheden tussen 2D en MPR beeldinterpretatie.

Met het digitale toetsprogramma VQuest is het mogelijk om zowel 2D als MPR beelden te gebruiken bij de toetsing van beeldinterpretatievaardigheden. In deze paragraaf wordt gerapporteerd over een onderzoek waarin is onderzocht welke kennis en vaardigheden vooral aangesproken worden bij 2D beeldinterpretatie en welke vooral bij MPR beeldinterpretatie.

2.2.2 Onderzoeksmethoden

Het onderzoek vond plaats onder twintig vierde- tot en met zesdejaars geneeskundestudenten die het coschap radiologie succesvol hadden afgerond. De deelnemers beoordeelden vier MPR casus en vier 2D casus uit de vier domeinen in het coschap (neuroradiologie, thoraxradiologie, abdominale radiologie en musculoskeletale radiologie) terwijl zij hardop nadachten. De uitgewerkte tekst werd afzonderlijk gecodeerd door twee onderzoekers met behulp van het codeerschema, gebaseerd op het raamwerk uit paragraaf 2.1. In geval van twijfel over de codering werd de andere onderzoeker geraadpleegd en werd consensus verkregen over de juiste code. Frequenties van de coderingen van de 2D vragen en de MPR vragen werden per deelnemer geproportioneerd en met elkaar vergeleken.

2.2.3 Resultaten

Uit de analyses blijkt dat het grootste deel van de uitingen van de deelnemers tijdens het hardopdenken werd gescoord als kennis of vaardigheid binnen de componenten 'perceptie' (41 %) en 'synthese' (28%). 'Analyse' had een minder groot aandeel (20%). De overige 11% bestond uit uitingen die niet onder beeldinterpretatie geschaard kon worden, bijvoorbeeld het oplezen van de vraag, opmerkingen over het gebruik van het toetsprogramma of metacognitieve opmerkingen. Bij de beoordeling van MPR beelden werden in totaal meer kennis en vaardigheid gescoord dan bij de beoordeling van 2D beelden. Hier staat tegenover dat de beoordeling van MPR beelden ongeveer twee keer zoveel tijd in beslag neemt.

Het aandeel van de componenten 'perceptie', 'analyse' en 'synthese' verschilt bij het interpreteren van 2D en MPR beeldvragen. De kennis en vaardigheden in de component 'perceptie' hebben het grootste aandeel bij de interpretatie MPR beelden en dit aandeel is significant groter dan bij de interpretatie van 2D beelden. Het aandeel van de component synthese is juist het grootst bij de interpretatie van 2D beelden en dit aandeel is significant groter dan bij de interpretatie van MPR beelden. Tot slot was het aandeel van het hardopdenken dat niet onder beeldinterpretatie geschaard kon worden in het geval van beeldinterpretatie van MPR beelden lager dan bij beeldinterpretatie in 2D. Het ging hierbij bijvoorbeeld om: het oplezen van de vraag, opmerkingen over het gebruik van het toetsprogramma of uitingen over het eigen zoekgedrag (metacognitief). Een groter deel van de denkprocessen werd dus besteed aan het interpreteren van het beeld.

2.2.4 Praktische bruikbaarheid

Aangezien de denkprocessen in alle componenten (perceptie, analyse en synthese) meer werden toegepast bij de beoordeling van MPR beelden ten opzichte van 2D beelden, zou geconcludeerd kunnen worden dat het gebruik van MPR beelden in alle gevallen beter zou zijn. Hier staat echter tegenover dat de beoordeling van MPR beelden meer tijd in beslag neemt; gemiddeld twee maal zoveel blijkt uit een ander onderzoek bij studenten en aijs, zonder hardopdenken (Ravesloot, conference proceedings RSNA 2013). Dit heeft consequenties voor het aantal toetsvragen dat in een toets kan worden opgenomen (hetgeen gevolgen kan hebben voor de betrouwbaarheid) of voor de duur van de toets (wat de praktische bruikbaarheid in de weg kan staan). Het is daarom van belang om goed te overwegen of het gebruik van het MPR beeld een meerwaarde heeft voor het beoogde toetsdoel van de vraag.

In dit kader is het interessant dat uit het hardopdenkonderzoek blijkt dat de 'perceptie' bij het beoordelen van een MPR beeld meer aangesproken wordt, terwijl bij een 2D beeld de nadruk ligt op synthese. Dit zou verklaard kunnen worden doordat in een 2D beeld de coupe met de afwijking reeds in beeld is, terwijl deze in een MPR beeld nog opgezocht moet worden. Ook zou een verklaring kunnen liggen in het feit dat een MPR beeld een grotere hoeveelheid aan visuele informatie bevat, zodat de perceptuele processen meer aangesproken worden. In de praktijk zou gekozen kunnen worden voor een MPR vraag om de nadruk te leggen op toetsing van 'perceptie', en juist voor een 2D vraag om de nadruk te leggen op toetsing van synthese. Voor het toetsen van analyse zou een vraagtype kunnen worden gebruikt dat analyse van de afwijking vergt, bijvoorbeeld door een meer-uit-meer vraag te gebruiken voor het beoordelen van de kenmerken van de afwijking.

2.3. Verschillen in kijk- en zoekgedrag tussen 2D en MPR beeldinterpretatie

2.3.1 Inleiding

Kijkgedrag vormt een belangrijke indicatie van hoe studenten beelden interpreteren. Dat geldt ook voor de wijze waarop studenten door beelden heen scrollen en van richting veranderen om tot een interpretatie te komen. Onderzoek naar hoe studenten naar beelden kijken en analyse van de digitale handelingen die ze daarbij verrichten is dan ook een noodzakelijke aanvulling op hardopdenkstudies (zoals gerapporteerd in de vorige paragraaf 2.2). Immers, in sommige gevallen zal een student hardop zeggen hoe hij of zij kijkt of wat zijn of haar zoekstrategie is, maar in de meeste gevallen gaat dat niet op. In deze paragraaf wordt onderzoek beschreven waarin kijkgedrag en zoekstrategieën tijdens beeldinterpretatie zijn bestudeerd met een oogbewegingsstudie en met een analyse van logfiles. Het verschil in kijk- en zoekgedrag bij de interpretatie van 2D en MPR beelden kan inzicht geven in de mate waarin perceptuele vaardigheden worden ingezet in deze vormen van beeldinterpretatie.

2.3.2 Onderzoeksmethoden

Procedure

In de studie werden tien tweedejaars studenten geïncludeerd. Zij kregen allen een radiologische anatomietoets met twintig aanwijsvragen, waarvan tien met 2D beelden en tien met MPR beelden. De vragen betroffen anatomische structuren in de borst- en buikholte, vergelijkbaar aan de vragen die de studenten drie maanden daarvoor hadden gemaakt (zie de radiologietoets voor tweedejaars geneeskundestudenten beschreven in hoofdstuk 3). Tijdens het maken van de toets werden de oogbewegingen en de digitale handelingen van de studenten geregistreerd. De digitale handelingen (scrollbewegingen en wisselingen in richting) waren nodig om te bepalen welke doorsnede van de scan door de student bekeken werd. De oogbewegingen en de geregistreerde handelingen werden aan elkaar gekoppeld. Per student en per vraag werden verschillende parameters berekend: de tijd tot aan de eerste fixatie van de relevante structuur (bijvoorbeeld de gevraagde anatomische structuur), het aantal fixaties in de relevante en irrelevante structuren en de totale fixatietijd in de relevante en irrelevante structuren. De relevante structuur was de structuur die gemarkeerd moest worden door de student

en was gedefinieerd als het antwoordmodel (de anatomische structuur was in 3D gemarkeerd in het beeld). Ook werden twee ruimtelijk inzichttests afgenomen. De eerste was een test op papier, waarbij identieke ruimtelijke figuren, die vanuit verschillende richtingen getekend waren, bij elkaar gezocht moesten worden (Mental Rotation Test, Peters 1995). De tweede was een gemodificeerde versie van de Visual Form Reconstruction Test (Smoker 1984). In deze test moesten drie legoconstructies nabouwd worden aan de hand van zijaanzichten op papier.

Deelnemers

De data van acht van de tien deelnemers waren bruikbaar voor analyses. Bij één deelnemer bleken de geregistreerde handelingen niet toereikend om de koppeling met de eye-tracking parameters te maken en bij één deelnemer bleek de registratie van de oogbewegingen onvoldoende. Verder kon een aantal vragen van de overige acht deelnemers niet gebruikt worden, in sommige gevallen omdat de oogbewegingen hier tijdelijk niet goed geregistreerd waren en in sommige gevallen omdat de deelnemer bij die vraag gebruik had gemaakt van de oriëntatiewisselingen waardoor een juiste koppeling aan de eye-tracking parameters niet mogelijk bleek. Om het aantal 2D en MPR vragen gelijk te houden werden vergelijkbare vragen van de andere vraagsoort geëlimineerd.

2.3.3 Resultaten

Uit de analyses bleek dat de studenten significant meer tijd nodig hadden om een relevante structuur te fixeren in een MPR beeld ten opzichte van een 2D beeld. Ook besteedden zij significant meer tijd en meer fixaties op de irrelevante structuur in MPR beelden ten opzichte van 2D beelden. Op de overige parameters deden zich geen verschillen voor. Verder bleek dat de scores op de ruimtelijk inzicht testen (met uitzondering van één van de drie lego-subtests) significant gerelateerd waren aan de score op de MPR beeldvragen, terwijl geen van de ruimtelijk inzicht testen significant gerelateerd was aan de scores op de 2D beeldvragen.

In een onderzoek bij dezelfde deelnemers naar het effect van MPR beelden op het de cognitive load (of working memory load) (Stuijzand et al. submitted) bleek bovendien bij extra analyses dat de tijd tot fixatie op een relevante structuur een significante voorspeller is van het goed of fout beantwoorden van een vraag.

2.4. Ontwikkeling van vaardigheden in radiologische beeldinterpretatie

2.4.1 Inleiding

Hoewel er nog weinig eenduidigheid bestaat over hoe vaardigheden in het interpreteren van radiologische beelden zich ontwikkelen kan op basis van de literatuur een aantal gemeenschappelijke kenmerken worden geformuleerd (Taylor 2007; Norman 1992). Taylor et al. geeft in zijn artikel (2007), een samenvatting van de tot dan toe gedane onderzoeken. Zo blijken in hardopdenkonderzoeken experts meer afwijkingen, kenmerken en diagnoses te benoemen dan novieten (Lesgold 1988). Hieruit zou geconcludeerd kunnen worden dat tijdens de ontwikkeling van beeldinterpretatievaardigheden zowel meer afwijkingen worden gepercipieerd (ontdekt), als ook dat er meer cognitieve vaardigheden en kennis ontstaan om afwijkingen te karakteriseren en diagnoses te genereren. Daarnaast blijkt uit een onderzoek waarbij experts en novieten in de radiologie werden gevraagd of ze een foto eerder hadden gezien, dat experts beter waren in het zich herinneren van afwijkende foto's en slechter in het zich herinneren van normale foto's dan novieten (Myles 1988). Dit zou erop kunnen wijzen dat radiologen, doordat ze veel variaties van een normaal beeld zonder afwijkingen hebben gezien, deze makkelijk herkennen en labelen als een "normaal beeld" en daarbij niet meer onthouden hoe dit beeld (deze variatie van normaal) er precies uitzag. Voor de klinische praktijk is het verschil tussen twee normale beelden immers irrelevant. Novieten daarentegen zouden deze conclusies mogelijk minder snel en goed kunnen trekken en de normale beelden als unieke exemplaren onthouden.

Een andere invalshoek bestrijkt onderzoek naar het effect van mate van ervaring van radiologen

op de sensitiviteit en specificiteit van radiologen voor het diagnosticeren van een bepaalde afwijking in een beeld. De resultaten van deze onderzoeken zijn niet consistent. Doorgaans wordt gevonden dat de accuratesse van de interpretatie toeneemt, vooral in studies waarin radiologen op verschillende expertiseniveaus worden vergeleken (studenten, artsen in opleiding tot radioloog, fellow radiologie, radiologen) (Nodine 1999). Wanneer een vergelijking wordt gemaakt binnen een expertiseniveau, bijvoorbeeld op basis van het aantal jaren ervaring of het aantal verrichtingen, zijn de resultaten wisselender (Taylor 2007). In sommige studies wordt gevonden dat de nauwkeurigheid van beeldinterpretatie afneemt met de jaren en dat net klare radiologen een hogere gevoeligheid hebben voor afwijkingen, bijvoorbeeld bij het opsporen van borstkanker op het mammogram (borstfoto) (Beam 2003, Barlow 2004), terwijl uit andere onderzoeken blijkt dat meer ervaren radiologen daarentegen weer minder vaak ten onrechte een afwijking diagnosticeren (Smith 2005). Opvallend is dat er voornamelijk onderzoek is verricht naar expertise bij de beoordeling van mammogrammen, hetgeen 2D beeldinterpretatie betreft.

Bij Nederlandse artsen in opleiding tot radioloog wordt halfjaarlijks een voortgangstoets afgenomen waarbij een deel van de vragen beeldinterpretatie toetst. Daarmee kan de ontwikkeling van beeldinterpretatie in de tijd worden gevolgd. In deze paragraaf wordt een onderzoek naar de kwaliteit van de voortgangstoets beschreven dat inzicht geeft in het verloop van de ontwikkeling van beeldinterpretatievaardigheden op basis van de uitkomsten van voortgangstoetsing in de radiologie (Ravesloot et al. 2012).

2.4.2 Methoden

Deelnemers

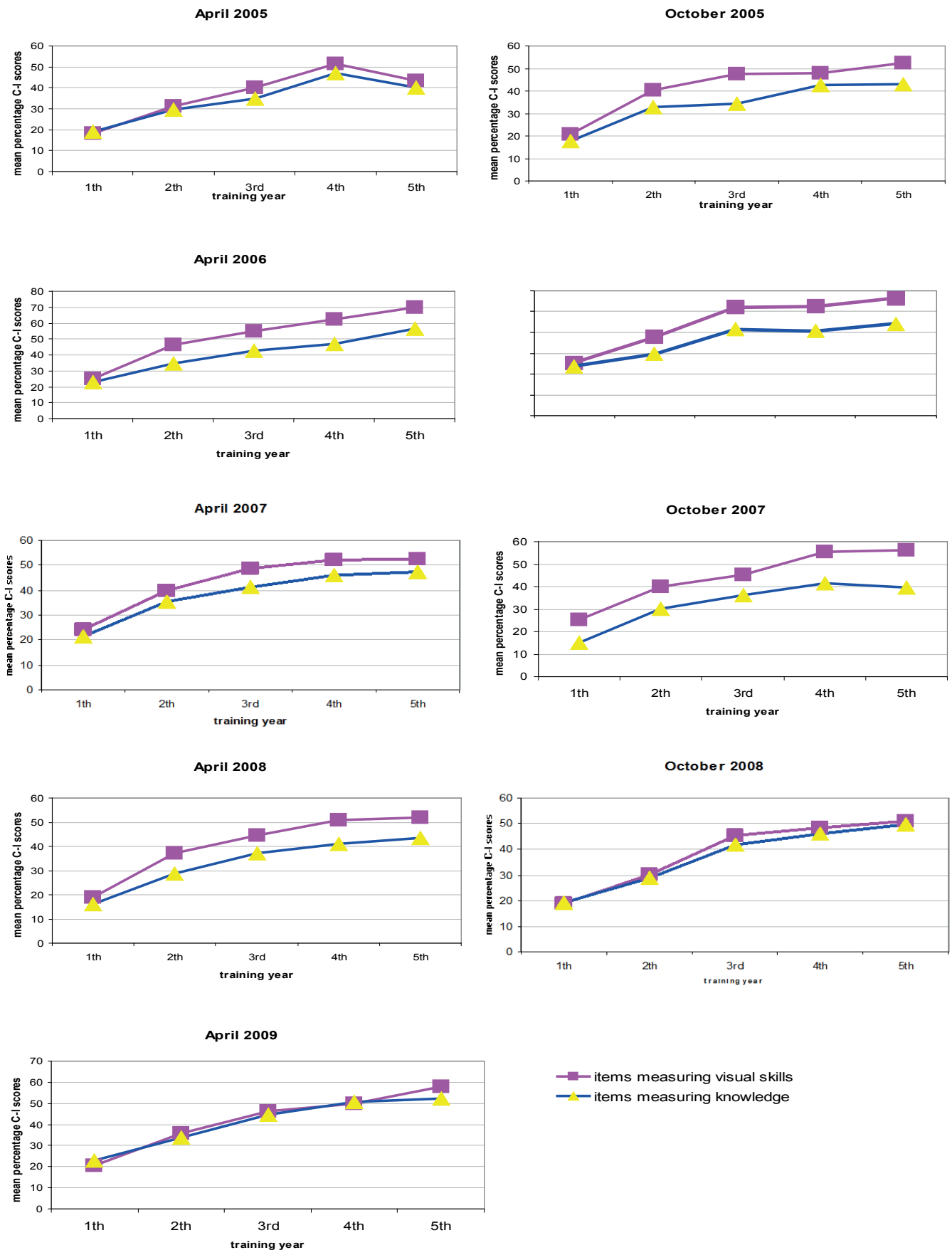
De radiologische voortgangstoets (VGT) wordt sinds 2003 halfjaarlijks afgenomen bij alle artsen in opleiding tot radioloog in Nederland (aios radiologie). Het niveau van de toets is gericht op het eindniveau van de vijfjaarlijkse opleiding. Alle aios maken de VGT halfjaarlijks gedurende de gehele opleiding. De aios volgen hun opleiding verspreid over 26 opleidingsziekenhuizen in Nederland. Voor de studie zijn alle toetsresultaten van de deelnemers aan de toetsen van 2005 tot 2009 geïnccludeerd. Het deelnemersaantal per toets lag gemiddeld rond 250 aios. In totaal werden van 498 aios 2281 toetsresultaten geanalyseerd.

Instrument

De toets bestaat uit 200 juist/onjuist/weet-niet vragen, waarbij ongeveer 30% van de vragen interpretatie van radiologische beelden betreft, beeldvragen, en de andere vragen alleen tekst bevatten, tekstvragen. Omdat het in die tijd nog een schriftelijke toets betrof werden er alleen 2D beelden in de toets gebruikt. Er werd een weet-niet optie toegevoegd, zodat deelnemers ervoor konden kiezen om een vraag niet te beantwoorden. Ze kregen voor de antwoordoptie 'weet niet' nul punten. Als ze een vraag wel beantwoordden kregen ze voor een goed antwoord één punt en voor een fout antwoord werd één punt afgetrokken. Dit wordt formula scoring genoemd en wordt onder meer gebruikt om gissen tegen te gaan, hetgeen de betrouwbaarheid van een toets positief zou beïnvloeden (McHarg 2005). De kans bestaat echter dat de keus voor de weet-niet optie meer door persoonlijkheid, gokbereidheid dan door kennis wordt bepaald (Muijtjens 1999).

Analyse

Bij de voortgangstoetsen met formula scoring werd een totaal score voor een toets berekend door het aantal foute antwoorden van het aantal goede antwoorden af te trekken en vervolgens het percentage van de maximale score (200 punten) te berekenen. Om een indicatie van de constructvaliditeit van de voortgangstoets te krijgen werden de toetsresultaten van negen opeenvolgende toetsen geëvalueerd (vanaf 2005 tot en met 2009) (Ravesloot 2012). Van elke toets werden de gemiddelde resultaten per opleidingsjaar met elkaar vergeleken (zowel voor de vragen met beelden als voor de vragen zonder beelden).



Figuur 2 Gemiddelde percentage goed-min-fout scores voor de tekstvragen (blauw) en de beeldvragen (roze) per opleidingsjaar per toets (VGT radiologie april 2005 t/m oktober 2009). Overgenomen uit het Ravesloot et al. (2012) "Construct validation of progress testing to measure knowledge and visual skills in radiology" *Medical Teacher*.

Eerst werd het gemiddelde van alle eerste-, tweede-, derde-, vierde en vijfdejaars berekend per toets en daarna is getoetst met ANOVA of er een significant verschil in scores is tussen de jaargroepen. Vervolgens is met post-hoc analyses nagegaan aan welke opleidingsjaren dit verschil is toe te schrijven.

2.4.3 Resultaten

Het aantal deelnemers per toets varieerde van 215 in april 2005 tot 284 deelnemers in april 2009. De betrouwbaarheid van de toetsen, geschat met Cronbachs alpha, ligt per toets rond de .90. Afzonderlijk voor de tekstvragen ($k_{\text{items}} \approx 160$) ligt dit rond de .85-.90; afzonderlijk voor de beeldvragen ($k_{\text{items}} \approx 30$) ligt dit rond de .60-.65. In figuur 2 zijn per toets de gemiddelde scores per jaargroep voor zowel de beeld- als de tekstvragen weergegeven.

Voor zowel de tekstvragen als de beeldvragen loopt de gemiddelde score op gedurende de opleiding. Deze stijging is het sterkst in de eerste drie jaren van de opleiding (Cohens effect-sizes variërend van .41 tot .55). Wanneer op groepsniveau het gemiddelde verloop van de score op de beeldvragen wordt vergeleken met de gemiddelde score op tekstvragen, blijkt dat de stijging voor de beeldvragen veel steiler verloopt en ook uiteindelijk hoger uitkomt aan het einde van de opleiding (zie voor een nadere toelichting van de analyses Ravesloot et al. 2012). Dit is in strijd met de aanvankelijke verwachting dat kennis en klinische ervaring de basis zijn voor beeldinterpretatievaardigheden en dat beeldinterpretatie pas later in de opleiding ontwikkelt na opbouwen van een grote kennis basis. Een kanttekening hierbij is dat het om cross-sectionele data gaat (vergelijkingen van opleidingsjaren per toets). Uit vervolgonderzoek blijkt dat ook longitudinaal dezelfde beeldvraagscore en tekstscorescurves worden gezien (Ravesloot, in voorbereiding 2014). Verder betrof het in deze voortgangstoetsen 2D beeldvragen. Deze vragen zouden onvoldoende representatief kunnen zijn voor beeldinterpretatie in de klinische praktijk.

2.4.4. Conclusie

Een belangrijke bevinding tijdens het bestuderen van de literatuur over beeldinterpretatie is dat er nog weinig bekend is over de ontwikkeling van kennis en vaardigheden die nodig zijn om beelden te kunnen interpreteren. Wel kunnen er verschillen tussen experts en novieten worden aangetoond. Bestaand onderzoek beslaat echter vooral 2D beelden en omdat het in de meeste studies een vergelijking tussen ervaren en minder ervaren radiologen betreft, zijn conclusies over hoe de vaardigheden zich ontwikkelen nog niet goed te trekken.

Tijdens een onderzoek naar de landelijke Nederlandse voortgangstoets radiologie (2005-2009) viel op dat de beeldvragen een sterkere stijging aan het begin van de opleiding kennen dan de kennis-score. Hoewel het ook hier alleen 2D beelden betreft, lijken beeldinterpretatievaardigheden zich vanaf het begin van de opleiding vrij snel te ontwikkelen. Dit resultaat dient in vervolgonderzoek verder te worden uitgezocht.

3. Optimalisering van de kwaliteit van toetsing van beeldinterpretatie

3.1. Inleiding

Een veronderstelling is dat bij digitale toetsing met MPR beelden, zoals toegepast in VQuest, kennis en vaardigheden voor beeldinterpretatie kwalitatief beter (in termen van validiteit en betrouwbaarheid) en efficiënter kunnen worden beoordeeld dan met 2D beelden (Ravesloot, submitted 2014-2). Er bestaat echter nog weinig evidentie voor de mate waarin en onder welke omstandigheden deze veronderstelling op gaat. In dit hoofdstuk zullen de resultaten van het onderzoek naar verschillen in toetskwaliteit tussen 2D en MPR beeldvraagtoetsen worden besproken (Ravesloot, RSNA 2013).

Over kwaliteitseisen die bij toetsing gelden is nationaal en internationaal veel gepubliceerd, zie bijvoorbeeld De Groot (1970), Eggen & Sanders (1993), Linn (1994), Linn, Baker & Dunbar (1991), Messick (1989, 1994, 1995), Stokking, Van der Schaaf, Erkens, & Jaspers (2004). Traditioneel worden aan de kwaliteit van toetsen twee eisen gesteld: validiteit en betrouwbaarheid. Deze eisen zijn gericht op het verkrijgen van meetresultaten die nauwkeurig zijn. Tegenwoordig worden aan beoordelingen van prestaties van studenten, waar voor hen diverse consequenties aan kunnen zijn verbonden, aanvullend ook eisen gesteld die zijn gericht op een juiste interpretatie en verantwoord gebruik van de resultaten, conform bekende en legitieme bedoelingen (consequentiële validiteit) (Messick 1989). Bovendien worden eisen van validiteit en betrouwbaarheid vaak aangevuld met de eis van praktische bruikbaarheid, wat inhoudt dat de toetsing haalbaar en efficiënt moet zijn (de toets moet bijvoorbeeld binnen een afzienbaar tijdsbestek kunnen worden afgenomen en er moeten geen onnodige kosten worden gemaakt). Deze eisen gelden ook voor het toetsen van beeldinterpretatievaardigheden. In dit hoofdstuk zullen de resultaten van een onderzoek naar de diverse aspecten van toetskwaliteit bij 2D en MPR toetsen worden besproken.

Validiteit

Validiteit is het belangrijkste criterium waar een toets aan moet voldoen en betreft de vraag of een toets meet wat ze beoogt te meten. De mate waarin digitale toetsing met MPR beelden kan bijdragen aan verhoging van de validiteit is nog ongewis. Te verwachten valt dat de authenticiteit van deze toetsvorm zou kunnen bijdragen aan een verhoging van de validiteit ten opzichte van traditionele toetsvormen. Immers, een authentieke vorm van toetsing doet idealiter meer recht aan de daadwerkelijke cognitieve processen die plaatsvinden tijdens het interpreteren van beelden en de beoordeling zal meer gestoeld zijn op het interpreteren van beelden in de praktijk. De verwachting is namelijk dat het herkennen en benoemen van afwijkingen en structuren in MPR beelden andere beeldinterpretatievaardigheden vereisen dan het herkennen en benoemen van dezelfde structuren op een enkele dwarsdoorsnede (2D beeld), zie ook hoofdstuk 2. Bij het interpreteren van radiologische MPR beelden, zoals CT-scans en MR-scans, zal de arts immers door alle dwarsdoorsneden heen moeten scrollen op zoek naar afwijkingen die duiden op een bepaalde ziekte. Bij een 2D beeld is de afwijking, indien aanwezig, in een oogopslag zonder scrollen te zien. Bij toetsing die bestaat uit 2D beelden wordt de deelnemer gevraagd aan de hand van slechts één of enkele naast elkaar geplaatste beelden van de honderden dwarsdoorsneden, van bijvoorbeeld een CT-scan, de juiste diagnose te stellen. Echter, in dit geval zal op deze 2D beelden de afwijking zichtbaar moeten zijn. Een belangrijke vaardigheid, namelijk het zoeken naar en herkennen van een afwijking (perceptie), wordt hierbij mogelijk minder valide getoetst.

Bovendien is een afwijking vrijwel altijd een 3D structuur. De derde dimensie kan in een 2D beeld niet (goed) worden beoordeeld, omdat aangrenzende coupes (vaak) ontbreken en de mogelijkheid ontbreekt om de afwijking in een andere doorsnedenrichting te bekijken. Dat is iets dat in de praktijk wel kan en als zodanig ook onderwezen wordt.

Betrouwbaarheid

Digitale toetsing kan standaardisering bij het nakijken van de toetsresultaten bevorderen, wat de betrouwbaarheid ten goede komt. Zo worden bijvoorbeeld bij het nakijken van aanwijsvragen in het pro-

gramma VQuest vooraf gedefinieerde overlays gebruikt om het antwoord volledig automatisch en digitaal te controleren. Bij een aanwijsvraag wordt de student gevraagd om een anatomische structuur of afwijking in het beeld aan te wijzen. Deze structuur is door de docent voorafgaand aan de toets (onzichtbaar voor de student) gemarkeerd. Dit wordt een overlay genoemd. Tijdens het nakijken wordt de overlay over de scan heen gelegd en wordt automatisch gecontroleerd of de door de student geplaatste marker binnen de overlay valt. Is dat het geval dan wordt het antwoord goed gerekend.

Daarnaast wordt er in VQuest ook gebruik gemaakt van longmenuvragen. Hierbij wordt gebruik gemaakt van een lange lijst van alternatieven, waar studenten het juiste antwoord uit moeten kiezen. De alternatievenlijsten kunnen voor de student onzichtbaar worden gemaakt, zodat het een open vraag benadert. Wanneer de student (ten minste) twee letters van het gekozen antwoord intypt, worden de alternatieven uit de lijst die deze twee letters bevatten, getoond. De student kan dan het gewenste antwoord aanklikken. Het nakijken van de vragen gebeurt gestandaardiseerd.

Praktische bruikbaarheid

Van belang is dat een toetsvorm, in dit geval toetsing met MPR beelden, praktisch bruikbaar is, zowel voor de docent als voor de student. Bovendien moeten de baten de kosten rechtvaardigen. Zo dient het toetsprogramma (VQuest) gebruiksvriendelijk te zijn en moet de technische uitvoerbaarheid van de toetsing acceptabel zijn. Belangrijk is ook dat de tijdsbesteding van studenten voor het maken van een MPR beeldvraag in vergelijking met een 2D beeldvraag haalbaar blijft, zodat er in een acceptabele toetsduur voldoende vragen kunnen worden afgenomen.

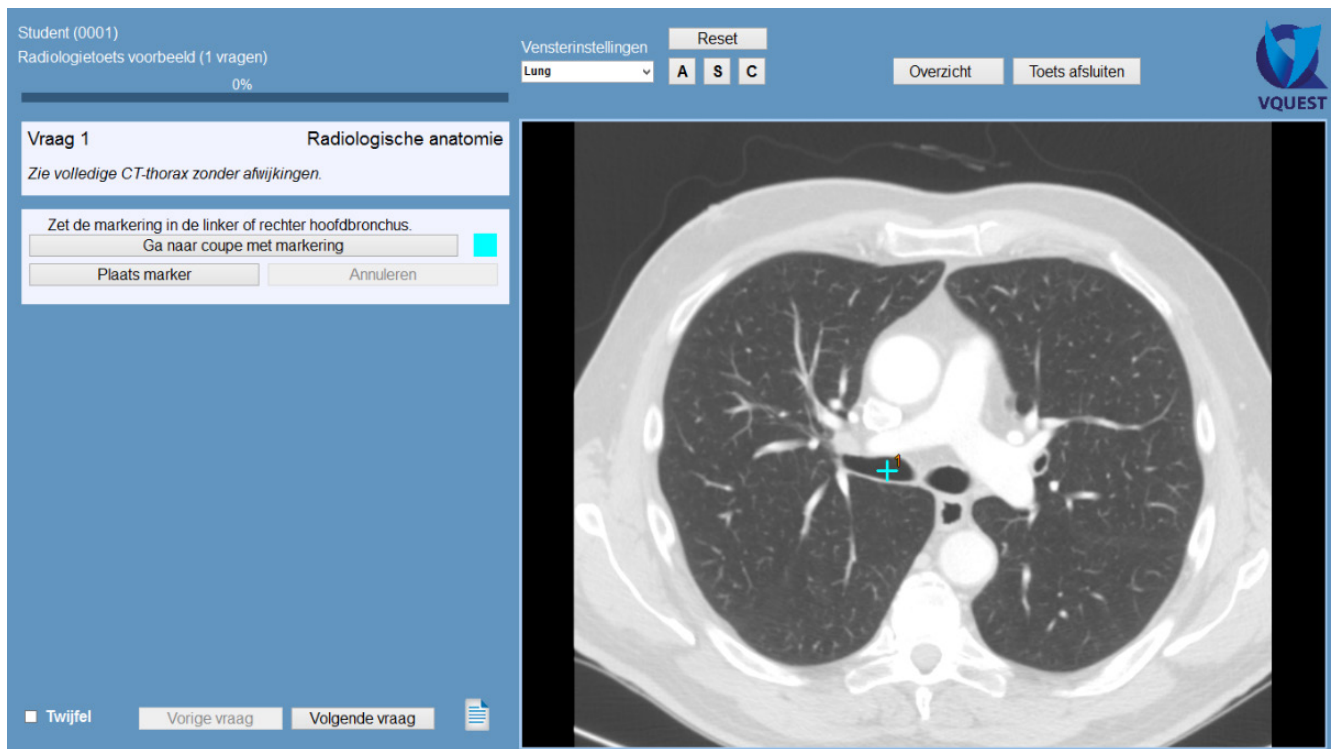
3.2 Onderzoeksmethoden

Onderzoekopzet en populatie

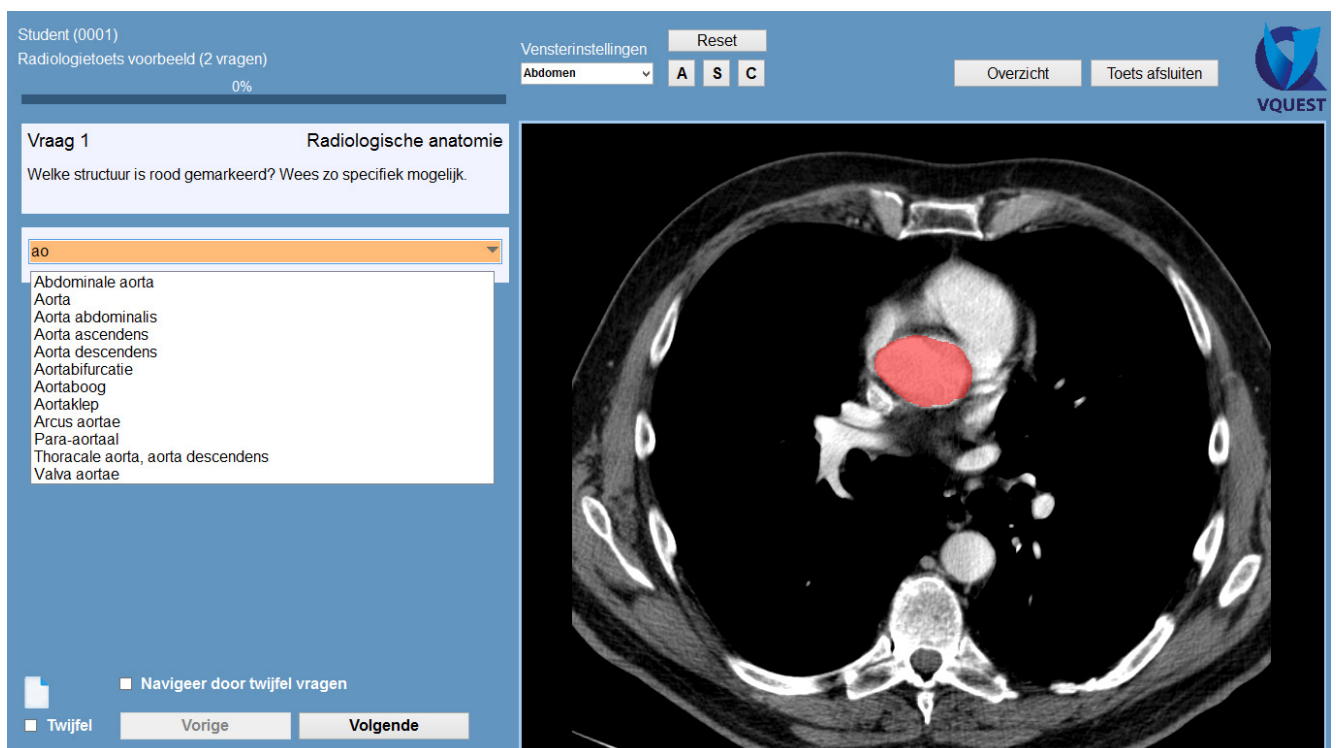
In het project zijn bij twee expertiseniveaus (tweedejaars geneeskundestudenten en vierde- tot en met zesdejaars coassistenten) toetsen met radiologische 2D en MPR beelden afgenomen. In 2012 is bij 246 tweedejaars geneeskundestudenten een radiologietoets afgenomen met twintig 2D beeldvragen en twintig MPR beeldvragen. De beeldvragen betroffen zowel aanwijs- als benoemvragen in een CT-scan van de borstkas (thorax) en de buik (abdomen). Bij een aanwijsvraag moest de student de gevraagde structuur in het beeld aanwijzen en bij een benoemvraag werd aan de student gevraagd de gemarkeerde structuur in het beeld te identificeren. Bij deze benoemvraag moest uit een lange lijst structuren (ruim 2000 anatomische structuren) de juiste worden gekozen (longmenuvraag). In figuur 3 is van beide vraagtypen een voorbeeld weergegeven (zie ook de appendix voor de studenthandleiding van VQuest). De 2D en de MPR beeldvraagtoets waren qua inhoud, vorm en niveau parallel aan elkaar, zodat gevonden verschillen in scores op toetsitems zoveel mogelijk aan het verschil in 2D of MPR konden worden toegeschreven.

Bij de tweedejaars geneeskundestudenten heeft een selectie van 33 studenten die de radiologietoets hadden gemaakt, ook een anatomietoets gedaan. Daarbij werd hun anatomiekennis getoetst aan de hand van een humaan stoffelijk overschot (snijzaaltoets). Deze snijzaaltoetsscore vormt de externe maat (gouden standaard) voor anatomiekennis en -vaardigheden.

Vanaf februari 2012 tot 1 september 2013 is bij 93 coassistenten radiologie een radiologietoets afgenomen ter afsluiting van hun coschap (klinische stage) radiologie. De toets bestond onder andere uit vijf 2D en vijf MPR beeldvragen. Het betrof zogenaamde casusvragen, waarbij aan de hand van een CT-scan met een afwijking enkele vragen werden gesteld. Vaak betrof dit een aanwijsvraag ("Wijs de afwijking aan"), een meer-uit-meer vraag waarbij de juiste kenmerken van de afwijking moesten worden aangegeven, en bijvoorbeeld een longmenuvraag waarbij de juiste diagnose gekozen moest worden. Daarnaast waren er inzicht- en kennisvragen over het gepresenteerde ziektebeeld (zie figuur 4 voor een voorbeeldvraag). Naast deze casusvragen over CT-scans, waren er ook beeldvragen over andere radiologische onderzoeken, bijvoorbeeld röntgenfoto's of echogrammen, en zaten er kennis- en inzichtvragen zonder beelden in de toets. De toets werd meestal elke 2 weken afgenomen in kleine groepjes, variërend van één tot acht coassistenten (zie figuur 5).

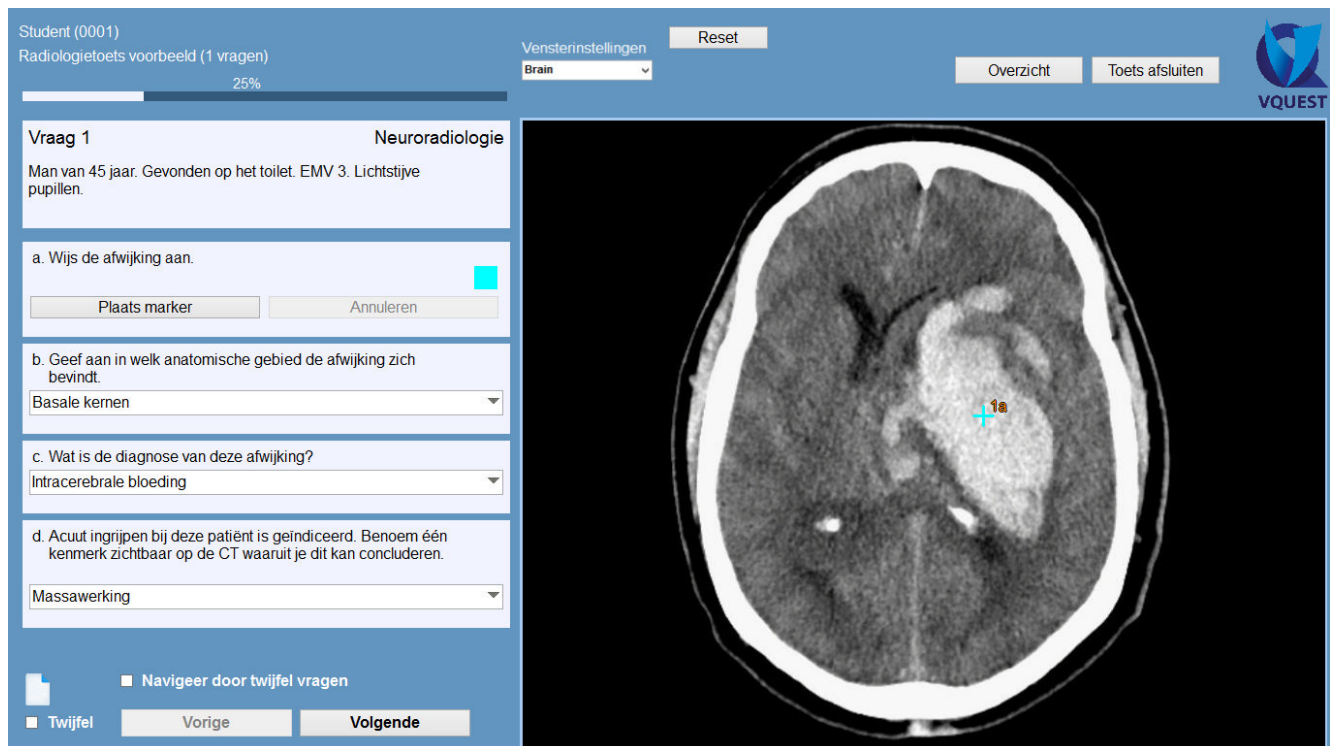


a



b

Figuur 3 Screenshot van aanwijsvraag (a) en benoemvraag (b) in radiologietoets bij tweedejaars geneeskunde studenten.



Figuur 4 Screenshot voorbeeld van een casusvraag coassistententoets. De gepresenteerde casus dient als voorbeeld en wordt niet gebruikt in de radiologietoets.



Figuur 5 Impressie tijdens afname cotoetsen

Analyses

Om een indicatie te krijgen van de ervaren authenticiteit, ervaren moeilijkheid en de gebruiksvriendelijkheid van de 2D en MPR beeldvragen, kregen de studenten na afloop van de toets een vragenlijst toegestuurd. Voor het verkrijgen van inzicht in de praktische bruikbaarheid zijn bovendien de tijdsbesteding van 2D en MPR beeldvragen vastgelegd. Daarnaast zijn er betrouwbaarheidanalyses van de toetsscores gedaan en zijn de scores op de twee beeldvraagtypen (2D en MPR beeldvragen) vergeleken. De scores op de 2D en MPR radiologische beeldvragen van de studenten werden gekoppeld aan de snijzaaltoetsscore. Hiermee werd een indicatie verkregen van de externe validiteit. De ervaring van docenten met de nieuwe toetsvorm is middels een korte vragenlijst en in onderwijsbijeenkomsten geëvalueerd.

3.3 Resultaten

Authenticiteit

Zowel de tweedejaars geneeskundestudenten als de coassistenten radiologie vonden de MPR beeldvragen beter aansluiten bij de klinische praktijk. Met name de coassistenten, die ook daadwerkelijk in de klinische praktijk stage lopen, konden dit goed inschatten. Bij de tweedejaars geneeskundestudenten werd daarom aanvullend gevraagd naar de aansluiting bij het genoten onderwijs. De studenten vonden ook dat de MPR beeldvragen hier beter bij aansloten.

Moeilijkheid

Een indicatie van de moeilijkheid van de vragen is op twee verschillende manieren verkregen. Ten eerste via het bevragen van de ervaren moeilijkheid in een vragenlijst aan studenten en ten tweede door p-waarden bij de toetsscores te berekenen.

Het resultaat is tweeledig. De studenten en coassistenten ervaren de MPR beeldvragen als makkelijker dan de 2D vragen, echter ze scoren niet hoger op de MPR beeldvragen. Sterker nog, in een aantal gevallen scoren ze zelfs significant lager op de MPR beeldvragen ten opzichte van de 2D vragen. Een mogelijke verklaring van deze bevinding zou kunnen zijn dat, doordat er meer informatie beschikbaar is in een MPR beeld (meer doorsnedes met mogelijkheid tot vervolgen van structuren), de student het gevoel krijgt dat het makkelijker is. Echter de taak op zich is mogelijk moeilijker, doordat de gevraagde structuur of afwijking niet meteen zichtbaar is in een MPR beeld, terwijl dit bij een 2D beeld wel het geval is. Een dergelijk fenomeen wordt bij openboektoetsen ook wel “false sense of security” genoemd (Dale 2009).

Uit het verschil in ervaren moeilijkheid kan worden geconcludeerd dat studenten het beantwoorden van MPR beeldvragen niet te moeilijk vinden. Dit is belangrijk voor de ervaren fairness ten aanzien van de toets. Docenten gaven ook aan dat het lastig is om één representatieve doorsnede uit een volumescan te selecteren voor het maken van een 2D beeldvraag zonder dat er onduidelijkheid bestaat over wat erop te zien is. De student kan immers niet scrollen of het contrast aanpassen in het 2D beeld. Aan de andere kant kent het gebruik van een MPR beeld ook uitdagingen, omdat er soms meerdere afwijkingen te zien zijn.

Betrouwbaarheid

De betrouwbaarheid geschat met Cronbachs alpha blijkt bij alle toetsafnames bij de tweedejaars geneeskundestudenten hoger uit te pakken voor de MPR beeldvragen dan voor de 2D beeldvragen. De spreiding van de scores op MPR beeldvragen blijkt groter te zijn dan die van de scores op 2D beeldvragen. Mogelijk komt dit door een betere differentiatie tussen studenten die over goede beeldinterpretatievaardigheden beschikken versus studenten die dit (nog) niet voldoende beheersen. Bij de coassistententoetsen kwam dit verschil in betrouwbaarheid niet eenduidig naar voren, maar mogelijk komt dit door het geringe aantal 2D en MPR beeldvragen in de toets.

Correlatie met een externe validatie maat (snijzaaltoetsscores)

Uit het onderzoek waarbij de beeldvraagsscores op de radiologietoets zijn vergeleken met de snijzaaltoetsscores, blijkt dat de MPR beeldscores goed met de snijzaaltoetsscores correleren. De 2D beeldvraagsscores correleren daarentegen niet met de snijzaaltoetsscores. Hoewel het hierbij om een kleine vrijwillige selectie van 33 studenten gaat is dit een positieve indicatie voor de validiteit van de MPR beeldvragen. Blijkbaar zijn de MPR beeldvragen, in tegenstelling tot 2D beeldvragen, redelijk goed in staat om de kennis van de 3D aspecten van de anatomie te meten.

Praktische bruikbaarheid en gebruiksvriendelijkheid toetsprogramma en beeldvragen

Voorwaardelijk voor het gebruik van MPR beeldvragen in digitale toetsing is dat het er niet toe zou moeten leiden dat het maken van de toets te ingewikkeld wordt voor de student en daarmee afleidt van het eigenlijke toetsdoel. Daarom zijn de studenten gevraagd naar hun mening over verschillende aspecten van de gebruiksvriendelijkheid van de gebruikte digitale toetstool en is de tijd die de studenten nodig hadden per MPR beeldvraag en 2D beeldvraag gemeten.

De studenten achtten de gebruiksvriendelijkheid van de MPR beeldvragen en 2D beeldvragen even hoog. Echter, de laadtijd van een MPR beeld vonden de studenten van de toets in april 2012 (te) lang. De laadtijd bleek ongeveer 10 seconden te zijn. Met een technische aanpassing in het toetsprogramma is de laadtijd gereduceerd naar 4 seconden, hetgeen door de studenten als acceptabel tot goed werd beschouwd. Sterke punten van de digitalisering van de toetsvragen vonden de studenten met name de verbeterde beeldkwaliteit en het gebruik van MPR beelden, waardoor studenten kunnen scrollen en ander beeldmanipulatie kunnen toepassen. Als nadeel werd ervaren dat studenten niet makkelijk konden zien welke vragen ze met twijfel hadden aangevinkt (als geheugensteuntje om eventueel later nog te bekijken). In het programma is daarom een extra functie ingebouwd die het nu mogelijk maakt om door de twijfelvragen te navigeren.

Bij de tijdsduurmetingen bij de toetsen van de tweedejaars geneeskundestudenten bleek dat de studenten gemiddeld twee keer zolang bezig zijn met het maken van een MPR beeldvraag (gemiddeld ongeveer 60 seconden) ten opzichte van een 2D beeldvraag (gemiddeld ongeveer 30 seconden). Hierbij is de tijd geteld van het moment waarop de student voor het eerst de vraag opent, totdat hij de vraag verlaat. Eventueel later terugkeren naar de vraag is hierbij dus niet meegenomen.

De docenten zijn over het algemeen positief over de nieuwe toetsvorm. Ze vinden dat deze beter aansluit bij de lesstof en de klinische praktijk. Bovendien vinden ze het geautomatiseerd nakijken een positief aspect. Wel kost het maken van de vragen en de antwoordmodellen van vooral de aanwijsvragen nog relatief veel tijd voor docenten.

Aanvullende bevindingen

Uit de resultaten van de toetsen voor coassistenten blijkt dat het gebruik van nieuwe vraagtypen, zoals longmenuvragen en aanwijsvragen, de docent meer inzicht geeft in de vaardigheden van de coassistent. Doordat bij een casus met deze vraagtypen de verschillende stappen van de beeldinterpretatie worden gevolgd (Van der Gijp 2014), kan worden vastgesteld waar het in het beeldinterpretatieproces fout gaat en wat voor soort fout het is (perceptiefout, analysefout, diagnosefout, kennishiaat). Bijvoorbeeld, soms bleek dat een coassistent niet de (juiste) afwijking had aangewezen in een beeld, maar uiteindelijk wel de juiste diagnose had gegeven bij de betreffende casus. Daaruit kon door de docent de conclusie worden getrokken dat de coassistent op basis van de bij de casus gegeven klinische informatie de juiste diagnose had gesteld (adequate kennis), maar het bij de perceptie al fout ging in het beeldinterpretatieproces en dus dat het gewenste beeldinterpretatieniveau niet gehaald werd. Andersom werd soms ook gezien dat een coassistent wel de juiste afwijking in het beeld had geïdentificeerd (juiste perceptie), maar vervolgens bij de benoeming van de kenmerken en de diagnosestelling de mist in ging.

3.4 Samenvatting

Het onderzoek bij geneeskunde studenten geeft indicaties dat het gebruik van MPR beeldvragen de kwaliteit van het toetsen van beeldinterpretatievaardigheden kan verhogen. Op vrijwel alle gemeten aspecten van toetskwaliteit scoorde de MPR beeldvraag hoger dan de 2D beeldvraag, zonder dat de praktische bruikbaarheid of gebruiksvriendelijkheid in het gedrang kwamen.

4. Stappenplan voor het afnemen van een digitale toets voor vaardigheden in radiologische beeldinterpretatie

4.1 Inleiding

In dit hoofdstuk worden de stappen voor het maken van een radiologische beeldinterpretatietoets besproken. Als bijlage is een studenthandleiding van VQuest toegevoegd. VQuest is het programma waarmee de beeldinterpretatietoetsen kunnen worden afgenomen. Een update van de handleiding en het stappenplan zijn ook op de website www.vquest.nl te vinden.

4.2 Toetsvragen maken en toets samenstellen

In deze paragraaf zullen in chronologische volgorde de stappen die genomen dienen te worden bij het opstellen van een toets worden besproken.

4.2.1 Toetsmatrijs

Voordat begonnen kan worden met het maken van vragen dient een toetsmatrijs te worden gemaakt. De toetsmatrijs moet aansluiten bij de geformuleerde leerdoelen in het onderwijs. Deze matrijs vormt een overzicht van de onderwerpen en vaardigheden die gedekt moeten worden in de toets (de 'content' van de toets). Ook wordt hierin aangegeven op welk niveau en met welke weging (hoeveel vragen moeten over dit onderwerp in de toets komen?) de onderwerpen en vaardigheden in de toets aan bod dienen te komen. Vaak wordt een toetsmatrijs samengesteld door een afvaardiging van vakinhoudelijke experts die over de toetsinhoud discussiëren en uiteindelijk tot een consensus komen.

Als voorbeeld zal hier kort de totstandkoming van de toetsmatrijs van de beeldinterpretatietoets voor coassistenten radiologie in het UMCU, zoals afgenomen tijdens het project Toetsing in Beeld (zie ook hoofdstuk4), worden besproken.

De leerdoelen van het coschap radiologie zijn zo vastgesteld dat net afgestudeerde basisartsen een goede radiologische basiskennis en beeldinterpretatie zullen bezitten van de meest voorkomende (sub)acute ziektebeelden. Het betreft het herkennen en diagnosticeren van (sub)acute ziektebeelden op diverse radiologische beelden, en kennis over welk radiologisch onderzoek zou moeten worden aangevraagd bij een (sub)acuut diagnostisch dilemma. Uit de lijst van ziektebeelden in de WHO-classificatie zijn alle ziektebeelden die aan deze criteria voldeden geselecteerd (WHO ICD-10,2007). Deze ziektebeelden zijn vervolgens onderverdeeld in traumatisch en niet-traumatisch. Daarna zijn de traumatische afwijkingen gerangschikt naar anatomische regio en de niet-traumatische ziektebeelden zijn onderverdeeld in neuroradiologie (ziekten van het zenuwstelsel), abdomen (ziekten van de buik), skelet en thorax (ziekten van de borstkas).

Tijdens enkele consensusbesprekingen, waaraan alle betrokken docenten van het coschap onderwijs in de radiologie van het UMCU deelnamen, is van elk ziektebeeld in deze lijst het niveau, waarop het ziektebeeld beheerst moet worden, bepaald. Hierbij is gebruik gemaakt van de herziene taxonomie van Bloom (Anderson & Krathwohl, 2001). Oorspronkelijk bestaat deze taxonomie uit zes niveaus in oplopende moeilijkheid van een cognitieve taak. Tijdens de consensusbesprekingen bleek voor het samenstellen van de leerdoelenlijst voor radiologische beeldinterpretatie voor coassistenten een vereenvoudigd model met drie niveaus geschikt. Niveau 1 (kennis) en 2 (begrip) uit het model zijn behouden en de oorspronkelijke niveaus 3 tot en met 6 zijn samengevoegd tot één niveau. Dit niveau (niveau 3 genoemd) omvat de beheersing van de daadwerkelijke beeldinterpretatievaardigheid. Niveau 1 betekent dat de coassistent kennis dient te hebben van het betreffende ziektebeeld (Wat is het? Bij wie komt het voor? Wat zijn de symptomen? Met welk radiologisch onderzoek zou je de ziekte kunnen aantonen?). Niveau 2 betekent dat de coassistent naast feitelijke kennis ook inzicht in het ziektebeeld moet hebben. Een voorbeeld van een dergelijk vraag is: Patiënt X heeft een inversietrauma van de enkel (naar binnen verzwikt) doorgemaakt. Welk type letsel van de enkel is hiermee geassocieerd? Bij niveau 3 dient de coassistent het ziektebeeld te kunnen herkennen op een radiologisch beeld. Bij een

ziektebeeld met niveau 3 wordt van de coassistent verwacht dat hij/zij aan de hand van een radiologisch beeld de juiste diagnose kan stellen, waarbij van hem/haar dus verwacht wordt alle fases en vaardigheden die komen kijken bij de beeldinterpretatie te beheersen. Naar de mening van de docenten is het in de praktijk namelijk niet zinvol om slechts één aspect van de beeldinterpretatie te beheersen; beeldinterpretatie is een complexe geïntegreerde vaardigheid.

Tot slot is vastgesteld hoeveel vragen er van elk van de onderdelen in een toets moeten zitten en wat een acceptabele toetsduur is.

4.2.2 Beelden selecteren

Voordat vragen kunnen worden geformuleerd, moeten eerst geschikte beelden worden verzameld. In de radiologie kan dit bijvoorbeeld worden gedaan door beelden te selecteren in de klinische praktijk. Bij het selecteren van MPR beelden is het van belang om op een aantal zaken te letten:

- 1) Zorg voor beschikking over de ruwe data (doorsnededikte < 1 mm). Anders is het in de toets niet mogelijk om in meer dan één richting door de dataset heen te scrollen. Bij zogenaamde gereconstrueerde beelden kan het beeld alleen in de gereconstrueerde richting worden beoordeeld.
- 2) Geadviseerd wordt om beelden te selecteren die alleen de beoogde afwijking bevatten. Bij meerdere afwijkingen bestaat namelijk het risico dat de student (bewust of onbewust) een andere dan de bedoelde afwijking kiest. Dat zou de validiteit van de vragen bij het beeld in gevaar kunnen brengen. Een suggestie is om beelden altijd door meerdere experts te laten screenen op eventuele secundaire afwijkingen. Wanneer er in het beeld naast de beoogde afwijkingen toch ook andere afwijkingen te zien zijn, kunnen de volgende aanwijzingen worden gehanteerd:
 - Probeer de niet-relevante afwijkingen alsnog kwijt te raken door de scan in te korten.
 - Geef, als bovenstaande niet mogelijk is, een duidelijke beschrijving van alle andere afwijkingen in de casustekst en geef duidelijk aan dat het in de vraag (of vragen) niet om de al beschreven afwijkingen gaat.
 - Formuleer de vragen en/of de casustekst zo dat alleen de bedoelde afwijking het juiste antwoord kan zijn.
- 3) Probeer de grootte van de scan te beperken (houd als stelregel maximaal 100 MB aan). Dat komt de laadtijd tijdens de toets ten goede.
- 4) Wees zorgvuldig met het anonimiseren van de beelden. Bij bijvoorbeeld Dicom-bestanden is aan iedere coupe patiëntinformatie gekoppeld!

4.2.3 Toetsvragen maken

Beeldbewerkingsfuncties

In de praktijk en ook in het toetsprogramma VQuest is het mogelijk om het beeld te bewerken. Zo kan bijvoorbeeld de grijswaardenverdeling (het contrast tussen verschillende weefsels in het beeld) worden aangepast en kunnen bij de MPR beelden de kijkrichting worden aangepast. Daarnaast is er ook nog een aantal geavanceerde beeldbewerkingsopties mogelijk, zoals bijvoorbeeld het aanpassen van de doorsnedendikte van het MPR beeld. Per toets of per beeldvraag kan worden ingesteld over welke beeldbewerkingsfuncties de student de beschikking heeft tijdens de toets. Hierbij is het van belang om de kosten-baten analyse goed te maken. Het toestaan van veel beeldbewerkingsfuncties kan de vraag soms nodeloos ingewikkeld maken, waardoor de kosten oplopen. Zo is een student vaak langer bezig met een dergelijke vraag (beeldbewerking kost tijd en zorgt voor meer visuele informatie), of kan door de vele beeldbewerkingsfuncties afgeleid worden ("Moet ik al die functies gebruiken? Dan zal het wel niet die overduidelijke in het oog springende afwijking zijn...."). Bij het bepalen van de beeldbewerkingsmogelijkheden kan het beste gedacht worden vanuit een toetsdoel. Welke relevante kennis en vaardigheden dient er te worden getoetst en welke beeldbewerkingsfuncties zijn daarbij nodig?

Casusbeschrijving

Bij een radiologische beeldvraag wordt het beeld vrijwel altijd begeleid door een casusbeschrijving. Dit is een korte tekst met bijvoorbeeld patiëntgegevens (geslacht, leeftijd, medische voorgeschiedenis) en

de klachtenpresentatie. Bij een radiologische anatomievraag (zoals in de eerder beschreven radiologie-toetsen voor tweedejaars geneeskundestudenten, zie hoofdstuk 3, kan een casustekst ook aangeven dat het om een beeld gaat zonder afwijkingen. In het algemeen geldt bij casusbeschrijvingen dat deze alleen relevante gegevens moeten bevatten. In het geval van beeldvragen kan door bewust informatie weg te laten de nadruk meer op beeldinterpretatie komen te liggen. Het gevaar bestaat namelijk dat de student door de klinische gegevens in de casusbeschrijving de juiste diagnose kan stellen zonder het beeld te hoeven interpreteren. Aan de andere kant is het voor de beantwoording van beeldvragen soms belangrijk om ook informatie over irrelevante andere afwijkingen in het beeld te vermelden.

Vraagtypen

Zoals eerder besproken zijn er naast de gebruikelijke vraagtypen, zoals meerkeuze- en meeruitmeervragen, ook meer geavanceerde vraagtypen beschikbaar die toegepast kunnen worden bij het toetsen van beeldinterpretatie. Het gaat hierbij om de aanwijsvraag en de longmenuvraag. Een aanwijsvraag kan bijvoorbeeld gebruikt worden om te toetsen of de student een afwijking of juist een normale anatomische structuur kan herkennen (perceptie). Een longmenuvraag benadert een open vraag en is daarom erg geschikt om te gebruiken bij het vragen naar de meest waarschijnlijke diagnose bij een beeld. Bij het bepalen van het vraagtype en het aantal vragen bij één beeld is het belangrijk om rekening te houden met het niveau van de deelnemers in relatie tot de moeilijkheidsgraad van het beeldmateriaal. Bij een duidelijke afwijking zal bijvoorbeeld een aanwijsvraag weinig toegevoegde waarde hebben (vrijwel iedereen zal hem goed hebben). In dat geval kan beter in de casusbeschrijving aangegeven worden welke afwijking te zien is (of gemarkeerd worden in het beeld) en naar de belangrijkste kenmerken en/of de diagnose gevraagd kunnen worden (diagnosevraag). Bij relatief makkelijk te interpreteren beelden zou gekozen kunnen worden voor een longmenuvraag in plaats van een meerkeuzevraag om het onderscheidend vermogen van de vraag te verhogen.

Itembank

Voor het samenstellen van een toets kunnen aan de hand van de toetsmatrijs steeds nieuwe vragen worden gemaakt (zie hierboven). Daarnaast is het mogelijk een itembank op te bouwen. Beelden en vragen kunnen zo opnieuw gebruikt worden. Dat geldt zowel voor complete vragen als voor onderdelen van de vraag. Zo kunnen beelden opnieuw gebruikt worden, met bijvoorbeeld andere vragen of een andere casusbeschrijving (beeldbank). Dat geldt ook voor de alternatievenlijsten bij de longmenuvragen. Deze alternatievenlijsten worden veel gebruikt in vragen die anatomische kennis toetsen en in diagnosevragen. Wanneer er één alternatievenlijst met anatomische structuren wordt gemaakt dan kan deze steeds voor alle anatomievragen worden gebruikt en indien nodig worden geüpdate. Hetzelfde geldt voor diagnoses. Voor de eerder besproken coassistententoetsen is per anatomische regio een diagnoselijst gemaakt, welke steeds wordt hergebruikt. Hergebruik van alternatievenlijsten bespaart veel tijd (alternatievenlijsten maken is daarentegen tijdrovend) en door hergebruik worden de lijsten na update vollediger en beter. Wanneer een nieuwe lijst voor het eerst wordt gebruikt is het aan te bevelen om studenten de mogelijkheid te geven om commentaar of missende alternatieven te laten noteren. Deze alternatieven kunnen vervolgens geëvalueerd worden om de lijsten te optimaliseren.

4.3 Toets afnemen

Voor het afnemen van een toets naar vaardigheid in beeldinterpretatie volgt hier een kort overzicht van de belangrijkste aandachtspunten en stappen. Deze punten zijn gebaseerd op de ervaringen die zijn opgedaan met het toetsprogramma VQuest.

4.3.1 Technische aspecten

Voorafgaand aan de toets

Omdat een beeldinterpretatietoets vaak vele (MPR) beelden bevat is de dataload van een dergelijk toets groot. Het is op dit moment dan ook niet mogelijk om de toets in VQuest online af te nemen. Dat

wil zeggen dat de toetsen altijd voorafgaand aan de toets op de betreffende computers wordt ingeladen. Afhankelijk van de snelheid van een eventuele draadloze netwerkverbinding, het aantal computers waarop getoetst gaat worden, en de grootte van de server, kan de distributie van de toetsen automatisch vanuit een centrale server naar de computers plaatsvinden. Om de installatie op de computers te laten starten moet handmatig worden ingelogd met een speciale inlogcode. Per computer kost het laden van de toets dan enkele minuten. In het geval van de toetsen bij de tweedejaars geneeskunde-studenten (het ging hierbij om ongeveer 150 computers) kostte dit distribueren ongeveer een half uur (30 computers konden tegelijkertijd de server benaderen om de toets te downloaden). Bij de planning van de toets moet men rekening houden met deze voorbereidingstijd.

Tijdens de toets

Voor het aanmaken van de inlogcodes maakt VQuest gebruik van een inlogfile. Hierin staan alle deelnemers zoals ingevoerd door de docent gekoppeld aan een uniek nummer (bijvoorbeeld studentID). Wanneer de toets begint vult de student zijn unieke nummer en het wachtwoord in. Bij summatieve toetsing wordt dit wachtwoord bij voorkeur pas vrijgegeven op het moment dat de toets begint. Het is niet mogelijk om met één nummer twee keer in te loggen. Dit voorkomt fraude. Er is een extra wachtwoord nodig om dan toch door te kunnen gaan (het extra wachtwoord is alleen bekend bij de surveillanten). Op deze manier kan een student met een probleem, bijvoorbeeld een gecrashte computer, toch doorwerken op een andere werkplek. De al ingevulde antwoorden worden door VQuest automatisch ingelezen van de server, waarna de student gewoon verder kan gaan met de toets. Het kan gebeuren dat een student deelneemt die niet officieel is aangemeld (bv door een administratieve fout) en daardoor niet in de inlogfile voorkomt. In dat geval is het mogelijk om een noodnummer te gebruiken. De surveillant moet dan een aantekening maken welke student bij het betreffende nummer hoort, zodat het toetsresultaat aan de juiste student gekoppeld kan worden.

Tijdens de toets worden de antwoorden van de student per vraag als een tekstbestandje weggeschreven naar een server en (optioneel) opgeslagen op de lokale harde schijf van de computer. Dit laatste dient als een back-up die na afloop van elke toets nog een keer gekopieerd wordt naar de server. Het antwoord wordt verstuurd naar de server op het moment dat de student de vraag afsluit door bijvoorbeeld op de “volgende vraag” button te klikken of door naar het vragenoverzicht te gaan. Wanneer de student later terugkeert naar de vraag en opnieuw de vraag verlaat, controleert het programma of er iets is gewijzigd in het antwoord en schrijft zo nodig een nieuw tekstbestandje met het veranderde antwoord naar de server. Dergelijke antwoordtekstbestanden zijn zeer klein (1kb) en belasten de server niet of nauwelijks. Het is dus zeer goed mogelijk om met grote aantallen studenten (meer dan 500) tegelijkertijd een toets af te nemen.

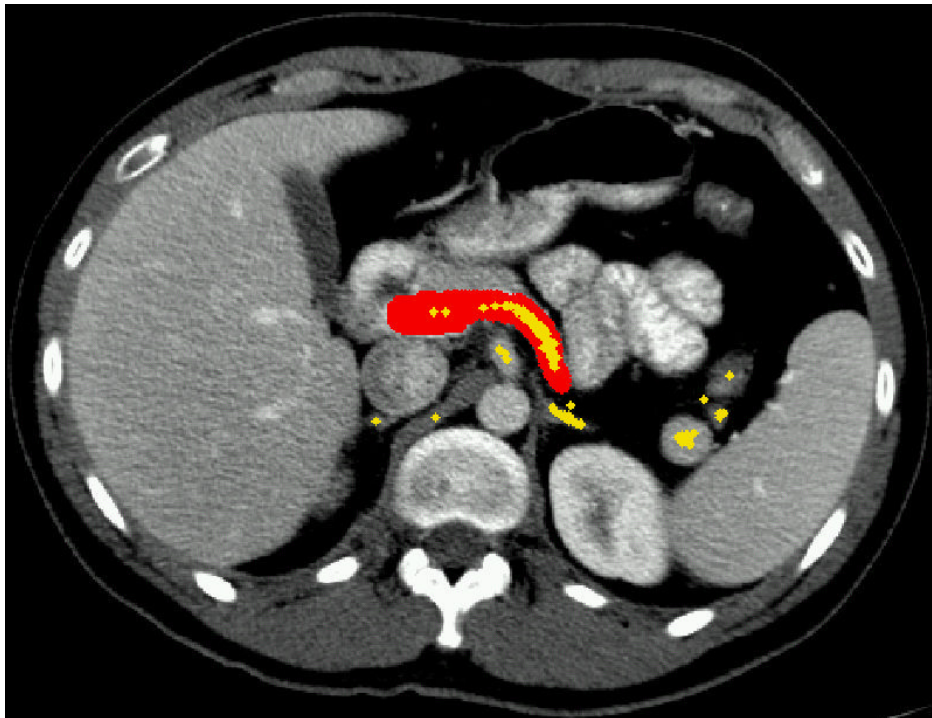
Na de toets

Na het afsluiten van de toets door de student wordt automatisch een script gestart dat de toets en het toetsprogramma van de computer verwijderd. De antwoorden van de studenten zijn immers op de server opgeslagen. Zo wordt ongewenste verspreiding van toetsvragen en toetsresultaten voorkomen.

4.3.2 Locatie

Software

Gebruik van 3D beelden in toetsen vergt een grotere opgave in databeheer dan toetsen met 2D beelden. De twee belangrijkste onderwerpen waarmee rekening gehouden moet worden zijn: opslaggrootte en inlaadsnelheid van de beelden. Opslaggrootte, omdat bij grote toetsen met veel 3D datasets veel data over het netwerk gestuurd moet worden van de server naar de toetscomputer voorafgaand aan de toets. Hierbij geldt hoe minder data, hoe minder voorbereidingstijd nodig is om de toets klaar te zetten. Inlaadsnelheid, omdat de wachttijden voor het laden van de beelden tijdens de toets zo klein mogelijk moet zijn. Bij het maken van keuzes voor bijvoorbeeld bestandsgrootte en compressietechnieken moet steeds een afweging worden gemaakt tussen de laadsnelheid tijdens de toets en de voorbereidingstijd voor de toets. Bijvoorbeeld, bij een bestandsformaat dat gebruikt maakt van een sterke compressie is



Figuur 6: Overzichtsafbeelding met alle gegeven antwoorden op een 2D beeld aanwijsvraag: Zet de marker in de vena lienalis (miltader). Het betreft een coupe uit een CT-abdomen. De gele stippen zijn de gegeven antwoorden van de studenten. Het rood gemarkeerde gebied betreft het door de docent ingekleurde antwoordmodel (vena lienalis in dit geval). Te zien is dat naast de vena lienalis diverse andere structuren gekozen zijn door de studenten, zoals bijvoorbeeld bijmilt, andere bloedvaten en darmstructuren. Een dergelijke overzichtsafbeelding van de antwoorden kan ook bij een MPR beeldvraag worden gegenereerd.

de voorbereidingstijd waarschijnlijk relatief kort, maar leidt dit tot een verslechtering van de laadsnelheid, omdat het bestand tijdens de toets uitgepakt moet worden. Een beeldtoets met VQuest kan in principe op iedere gangbare Windows-based computer worden afgenomen. Meestal draait hierop een 32-bits besturingssysteem waarbij beelden tot 100Mb zonder problemen kunnen worden weergegeven. Voor de coassistententoets in het project Toetsing in Beeld (hoofdstuk 3) is een speciale toetsruimte ingericht waarop een 64-bits Windows-versie draait. Daarbij kunnen ook beelden die groter dan 100Mb zijn zonder problemen worden weergegeven (zeker 500-700 Mb).

Toetsruimte

Belangrijke aandachtspunten bij het inrichten of selecteren van een toetsruimte voor een beeldinterpretatie toets zijn: lichtinval, reductie van afwijkmogelijkheden en stand van computerscherm. Uiteraard zijn deze aspecten bij alle digitale toetsen van belang, maar bij toetsing met beelden zijn ze fundamenteel.

Bij voorkeur dient de ruimte volledig te verduisteren te zijn met de mogelijkheid tot dimbaar kunstlicht. Zonlicht is namelijk niet te reguleren en is vaak zeer sterk. Hierdoor neemt de beeldkwaliteit op het computerscherm zeer sterk af, waardoor de validiteit van de toetsscores zal dalen. Wanneer zonlicht niet te weren is uit de beoogde toetsruimte, wordt sterk aanbevolen om na zonsondergang te toetsen. In het geval van de tweedejaars radiologietoetsen in het project Toetsing in Beeld is dit steeds gedaan.

De beelden in een beeldinterpretatietoets worden groot weergegeven op het beeldscherm. Daardoor wordt zoveel mogelijk voorkomen dat fouten door studenten het gevolg zijn van een te kleine weergave. Bovendien is de beeldweergave hierdoor zo vergelijkbaar mogelijk met de klinische praktijk, waarin de beelden beeldvullend worden weergegeven. Een nadeel hiervan is echter dat op vrij grote afstand vaak te herkennen is aan de hand van het weergegeven beeld bij welke vraag een student is.

De mogelijkheid tot afkijken zou daarom potentieel hoger kunnen zijn als hier met de plaatsing van de computers (en/of studenten) geen rekening wordt gehouden.

Het verdient de voorkeur om instelbare beeldschermen te gebruiken. Op die manier kan deze op de persoon worden aangepast en kan gecorrigeerd worden voor spiegeling door lichtinval.

4.4 Toets nakijken

4.4.1 Nakijken in VQuest

Na afloop van de toets staan de antwoorden van alle studenten op de server. Alle antwoorden van één student staan in één mapje. De antwoordbestanden kunnen in de docent-module van VQuest ingelezen worden en vervolgens automatisch worden nagekeken. Het resultaat is een matrix van studenten versus vragen, met per vraag één kolom met het gegeven antwoord en één kolom met de bijbehorende score. Daarmee heeft de docent een overzicht over welke antwoorden vaak gegeven zijn. Vooral bij longmenuvragen kan het zinvol zijn om na te gaan wat de variatie is in gegeven antwoorden.

4.4.2 Nakijken van aanwijsvraag

Voor het nakijken van een aanwijsvraag dient de docent een *overlay* als antwoordmodel te hebben toegevoegd. Een overlay is een gemarkeerd gebied in het beeld. Dit kan zowel een volume (bij MPR beelden) als een oppervlak (bij 2D beelden) zijn. Tijdens het nakijken wordt de overlay met de antwoorden van de student ingelezen in VQuest. Als de student de marker in het gemarkeerde gebied heeft gezet wordt het antwoord goed gerekend. Als ondersteuning bij het nakijken van een aanwijsvraag kunnen de antwoorden van alle deelnemers worden weergegeven in één overzichtsafbeelding (zie figuur 6). De docent creëert daarmee een overzicht van de distributie van de antwoorden en kan naar aanleiding hiervan overwegen om het antwoord model (de overlay) eventueel aan te passen. Af en toe blijkt dat namelijk noodzakelijk. Een voorbeeld uit het project Toetsing in Beeld betreft een toetsvraag met een 2D beeld van een CT-scan van de longen (sagittale coupe rechter long CT-thorax). Daarbij werd de studenten gevraagd om de fissura horizontalis (dit is een spleet tussen de rechter bovenkwab en de middenkwab) aan te wijzen. Bij het nakijken van de toetsen bleek dat een aantal studenten de marker strategisch op het kruispunt van twee fissuren had gezet. Op de antwoordoverlay was de fissura minor in zijn geheel (dus ook de kruising met de andere fissuur, de fissura obliqua) ingekleurd en werd het antwoord van deze studenten (ten onrechte) volledig goed gerekend. Op basis van deze bevinding hebben de docenten de overlay achteraf aangepast.

4.4.3 Nakijken longmenuvraag

Bij een longmenuvraag heeft de docent vooraf aangegeven welk alternatief (of welke alternatieven) goed moeten worden gerekend. Tijdens het nakijkproces wordt gecheckt of het door de student gegeven antwoord in de lijst van juiste antwoorden voorkomt. Vaak zijn meerdere alternatieven uit de longmenu juist, omdat in het medische jargon vele synoniemen kent. Ook worden er vaak deelpunten aan alternatieven toegekend. Het komt namelijk regelmatig voor dat een antwoord wel juist is, maar niet zo precies. Bijvoorbeeld, bij een fractuur (breuk) van processus spinosus van C7 (botuitsteeksel van halswervel 7) wordt gevraagd waar de afwijking zich bevindt. Het beste antwoord is in dit geval processus spinosus van C7, maar een student kan bijvoorbeeld ook alleen maar “C7” of zelfs alleen “nekfractuur”. Aan deze laatste 2 alternatieven zouden dan minder of zelfs geen punten kunnen worden toegekend. Het is noodzakelijk om in de vraag steeds aan te geven dat de student “zo specifiek mogelijk” dient te zijn. Hiermee worden discussies achteraf vermeden.

5. Conclusies

Uit de besproken onderzoeken blijkt dat MPR beeldvragen positief kunnen bijdragen aan de kwaliteit van toetsing van radiologische beeldinterpretatie. In vergelijking met 2D beeldvragen werden MPR beeldvragen als authentieker ervaren, leveren ze vaak betrouwbaarder toetsscores op en correleren ze beter met een externe validatiemaat. Bovendien blijkt uit de hardopdenkstudies dat bij de interpretatie van MPR beelden meer en deels andere vaardigheden moeten worden ingezet dan bij de interpretatie van 2D beelden in hetzelfde domein. Mogelijk vormt dit een deel van de verklaring voor de bevinding dat de toetsscores van MPR beeldvragen meestal lager zijn. Daarnaast blijkt toetsen met MPR beelden praktisch haalbaar. Dat betreft zowel de fase van het maken van de toetsvragen, als de fase van het (technisch en logistiek) afnemen van de toetsen.

6. Referentielijst

- Anderson, LW, & Krathwohl DR (eds.) (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman
- Azevedo R, Lajoie S, Desaulniers M, Fleiszer D, Bret P. Radtutor: the theoretical and empirical basis for the design of a mammography interpretation tutor. In: du Boulay B, Mizoguchi R, editors. *Artificial Intelligence in Education*. Amsterdam: IOS Press; 1997.
- Beutel J. *Handbook of medical imaging: Physics and psychophysics*. Washington: SPIE; 2000.
- Beam CA, Conant EF, Sickles EA. Association of volume and volume independent factors with accuracy in screening mammogram interpretation. *J Natl Cancer Inst* 2003; 95:282–290.
- Barlow WE, Chi C, Carney PA, et al. Accuracy of screening mammography interpretation by characteristics of radiologists. *J Natl Cancer Inst* 2004; 96:1840–1850.
- Cooper L, Gale A, Darker I, Toms A, Saada J. Radiology image perception and observer performance: how does expertise and clinical information alter interpretation? Stroke detection explored through eye-tracking. Paper presented at Proceedings of SPIE: Medical Imaging, February 7, 2009, Lake Buena Vista, FL, USA.
- Dale, VH, Wieland, B, Pirkelbauer, B et al. Value and benefits of open-book examinations as assessment for deep learning in a post-graduate animal health course. 2009; 36:403-10.
- De Groot, A.D., Some badly needed non-statistical concepts in applied psychometrics. *Nederlands Tijdschrift voor de psychologie* 1970; 25, 360-376.
- Eggen, TJHM., & Sanders, PF (1993), *Psychometrie in de praktijk*. Arnhem: Cito.
- Hegarty M, Components of Spatial Intelligence. *Psychology of Learning and Motivation*, 2010; Vol 52, pp.265-297
- Krupinski EA. Visual scanning patterns of radiologists searching mammograms. *Acad Radiol*. 1996 Feb;3(2):137-44.
- Kundel HL, La Follette PS, Jr. Visual search patterns and experience with radiological images. *Radiology*. 1972 Jun;103(3):523-8.
- Kundel HL, Nodine CF, Carmody D. Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Invest Radiol*. 1978 May-Jun;13(3):175-81.
- Kundel HL, Nodine CF, Conant EF, Weinstein SP. Holistic component of image perception in mammogram interpretation: gaze-tracking study. *Radiology*. 2007 Feb;242(2):396-402.
- Lesgold AM, Robinson H, Feltovich P, Glaser R, Klopfer D, Wang Y. Expertise in a complex skill: diagnosing x-ray pictures. Chi MTH, Glaser R, Farr MJ, editors. Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1988.
- Linn, R.L., Performance assessment. Policy promises and technical measurement standards. *Educational Researcher* 1994; 23 (9), 4-14.
- Linn, R.L., Baker, E.L., & Dunbar, S.B., Complex, performance-based assessment: expectations and validation criteria. *Educational Researcher*, 1991; 20 (8), 15-21.
- McHarg, J, Bradley, P, Chamberlain, S et al. Assessment of progress tests. 2005; 39:221-7.
- Messick, S., Validity. In R.L. Linn (Ed.), *Educational Measurement* 3rd ed. 1989; pp. 13-103. New York: Macmillan.
- Messick, S., The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher* 1994; 23 (2), 13-23.
- Messick, S., Validity of psychological assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist* 1995; 50 (9), 741-749.
- Morita W, Miwa K, Kitasaka T, Mori K, Suenaga Y, Iwano S, et al. Interaction of perceptual and conceptual processing: expertise in medical image diagnosis. *International journal of human-computer studies*. 2008; 66(5): 370-90.
- Muijtjens, AM, Mameren, HV, Hoogenboom, RJ et al. The effect of a 'don't know' option on test scores: number-right and formula scoring compared. 1999; 33:267-75.
- Myles-Worsley, M, Johnston, WA, Simons, MA. The influence of expertise on X-ray image processing. *American Psychological Association* 1988; 14:553-7.
- Nodine CF, Kundel HL. The cognitive side of visual search. O'Regan JK, Lévy-Schoen A, editors. North-Holland: Elsevier Science Publishers B.V.; 1987.

- Nodine CF, Kundel HL, Mello-Thoms C, et al. How experience and training influence mammography expertise. *Acad Radiol* 1999; 6: 575–85.
- Nodine CF, Mello-Thoms C, Kundel HL, Weinstein SP. Time course of perception and decision making during mammographic interpretation. *AJR Am J Roentgenol*. 2002 Oct;179(4):917-23.
- Norman, GR, Coblenz, CL, Brooks, LR et al. Expertise in visual diagnosis: a review of the literature. 1992; 67:S78-S83.
- Peters M, Laeng B, Latham K, Jackson M, Zaiyouna R, Richardson C. A redrawn vanderberg and kuse mental rotations test: different versions and factors that affect performance. *Brain and cognition*. 1995(28):39-58.
- Raufaste E, Eyrolle H, Mariné C. Pertinence generation in radiological diagnosis: spreading activation and the nature of expertise. *Cognitive Science*. 1998;22:517-46.
- Raveslout, CJ, Van der Gijp, A, Van der Schaaf, MF, Ten Cate, ThJ, Van Schaik, JPJ, Tipker, CA, Maas, M, Vincken, KL, Quality improvement of radiological image interpretation skills assessment through digital MPR multi-slice images in medical education, conference proceedings RSNA 2013
- Raveslout, CJ, van der Schaaf, MF, Haaring, C Kruitwagen, CLJJ, Beek FJA, Ten Cate, ThJ, Van Schaik, JPJ Construct validation of progress testing to measure knowledge and visual skills in radiology. *Medical Teacher* 2012; 34:1047-55.
- Raveslout, CJ, Van der Schaaf, MF, Kruitwagen, CLJJ, Haaring C, Ten Cate, ThJ, Predictors of expertise development in radiology (working title), in preparation (2014-1).
- Raveslout, CJ, Van der Schaaf, MF, Van Schaik, JPJ, Ten Cate, ThJ, Van der Gijp, A, Mol, CP, Vincken KL, Improved Testing of Radiological Image Interpretation Skills using Volumetric Images, under revision (2014-2).
- Rogers E. A study of visual reasoning in medical diagnosis. Cottrell GW, editor. Mahwah, N.J.: Lawrence Erlbaum; 1996.
- Smith-Bindman R, Chu P, Miglioretti DL, et al. Physician predictors of mammographic accuracy. *J Natl Cancer Inst* 2005; 97:358–367.
- Smoker WR, Berbaum KS, Luebke NH, Jacoby CG. Spatial perception testing in diagnostic radiology. *AJR Am J Roentgenol*. 1984 Nov;143(5):1105-9.
- Stokking, K.M., Van der Schaaf, M.F., Jaspers, J., & Erkens, G., Teachers' assessment of students' research skills. *British Educational Research Journal* 2004; 30 (1), 93-116.
- Stuijzand B.G, Van der Schaaf MF, Kirschner FC, Raveslout CJ, Van der Gijp A, Vincken KL (submitted), Medical Students' Working Memory Load in Multiplanar Image Interpretation: Insights from Human-Computer Interactions and Eye Movements.
- Taylor, PM. A review of research into the development of radiologic expertise: implications for computer-based training. 2007; 14:1252-63.
- Van der Gijp, A, Van der Schaaf, MF, Van der Schaaf, IC, Huige, JCBM, Raveslout, CJ, Van Schaik, JPJ, Ten Cate, ThJ , 'Interpretation of radiological images: towards a framework of knowledge and skills'. *Advances in Health Sciences Education*. (2014)
- Vincken, KL, Mol, CP & Raveslout, C.J. (2011). VQuest.; www.vquest.nl
- WHO ICD-10 (version 2007), (<http://apps.who.int/classifications/apps/icd/icd10online/>)

Appendix: Student Handleiding VQuest



Inleiding

Met dit digitale toetsprogramma kunnen – naast vragen aan de hand van één enkele coupe uit een CT-scan – ook vragen aan de hand van gehele CT-scans worden gesteld. Bij een vraag over een volledige CT-scan/MRI-scan is het mogelijk om door de scan te scrollen en soms in meerdere richtingen te bekijken. Hieronder volgt de gebruikshandleiding. Daarnaast is online een instructiefilm van het programma te bekijken (www.vquest.nl/Instructiefilm.html).

Algemeen

Inloggen

Gebruikersnaam: *studentnummer*

Wachtwoord: let op instructies van de surveillant

Voortgangsbalk

Linksboven in beeld is een voortgangsbalk zichtbaar. Deze geeft met een percentage bij benadering aan welk deel van de vragen je al beantwoord hebt.

Volgende/Vorige vraag

Wanneer je klaar bent met het beantwoorden van een vraag, kan je met de button “Volgende vraag” naar de volgende vraag gaan (via “Vorige vraag” kun je dus ook naar de vorige vraag gaan).

Overzicht

Door op de knop “Overzicht” te klikken krijg je een overzicht van alle vragen in de toets. Deze zijn gerangschikt per onderwerp. Een beantwoorde vraag is blauw, de nog niet beantwoorde vragen zijn dikgedrukt en wit. Door de cursor boven een vraagnummer te houden, verschijnt de casustekst in beeld. Wanneer je op een vraagnummer klikt, ga je naar de desbetreffende vraag toe. Wanneer je op “Sluiten” klikt, blijf je bij de huidige vraag.

Twijfelbox

Onderaan iedere vraag is de mogelijkheid om “Twijfel” aan te vinken (linksonder in beeld). Dit is puur bedoeld als geheugensteuntje voor jezelf. Als je deze aanvinkt verschijnt onder het betreffende vraagnummer in het overzicht een vraagteken. Op die manier kun je vragen markeren, waar je later nog een keer naar wilt kijken. Door de optie “navigeer door twijfelvragen” linksonder in beeld aan te vinken, navigeer je met de “volgende..” en “vorige..” toetsen door alleen de “met-twijfel-gemarkeerde” vragen. **NB: het antwoord dat je hebt gegeven op deze vraag wordt altijd gewoon nagekeken en het aangevinkt houden van de twijfelbox heeft geen consequenties voor de puntentelling.**

Commentaarvak

Als je een opmerking over een vraag hebt, klik dan op het memo-icoontje linksonder in beeld. Er opent dan een pop-up venster met een commentaarveld. Hier kun je jouw opmerking intypen en opslaan.

Toets inleveren

Wanneer je klaar bent met de toets, klik je op de button “Toets afsluiten” boven in beeld. Je ziet dan in het vragenoverzicht welke vragen je nog niet beantwoord hebt en of je nog vraagtekens hebt staan. Door op een nummer te klikken kun je deze vraag bekijken en eventueel het antwoord wijzigen. **NB: het is niet nodig om de vraagtekens weg te halen. De gegeven antwoorden worden gewoon nagekeken.**

Bekijken en bewerken van beelden

Bekijken van volledige scans

Scrollen:

Dit kan op 2 manieren:

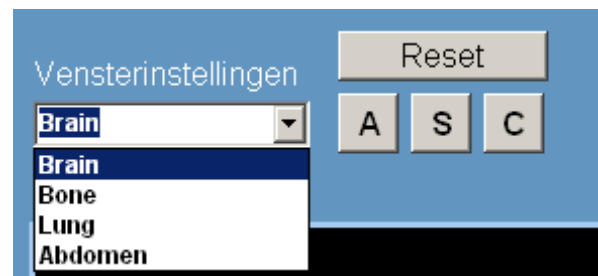
- 1 Door in de viewer je linker muisknop ingedrukt te houden en de muis van boven naar beneden te schuiven (je scrollt dan snel door de scan heen).
- 2 Door aan het muiswiel te draaien (je scrollt dan coupe naar coupe).

NB: Onderstaande beeldbewerkingsfuncties kunnen afhankelijk van de vraag aan of uit staan:

Wisselen van vensterinstelling

Dit kan op 2 manieren:

1. Met het dropdownmenu boven in het scherm kun je uit een aantal preset vensterinstellingen kiezen: *bone* = bot, *abdomen* = wekedelen en *lung* = long (figuur 1).
2. Handmatig kun je de vensterinstelling aanpassen door het scrollwiel van je muis ingedrukt te houden en van links naar rechts en/of van boven naar beneden over het beeld heen te bewegen.



Figuur 1

Wisselen van doorsnederichting

Dit kan met de knoppen **A** (axiaal), **C** (coronaal) en **S** (sagittaal) bovenaan het scherm (figuur 1).

Zoomen

Met de “plus” en “min” toetsen op het cijferblok rechts op het toetsenbord (numpad) kun je respectievelijk in- en uitzoomen.

Aanpassen van de rendering methode en coupedikte

Soms is het mogelijk om de renderingmethode aan te passen bij het bekijken van een scan. Er verschijnt dan rechts van de resetknop een dropdownmenu met de verschillende renderingopties. Dit werkt alleen bij een coupedikte groter dan 1. De coupedikte kan handmatig worden aangepast door in het venster onder het renderingmenu een getal in te typen.

Bekijken van 2D beelden

Bij één of enkele coupes uit een CT of MRI-scan of bij een conventionele of echo-opname kan je de grijswaardenverdeling handmatig wijzigen: muiswiel ingedrukt houden en over het beeldscherm bewegen. In het geval van een CT-coupe kan je voor het wijzigen van de vensterinstellingen ook het dropdownmenu preset vensterinstellingen gebruiken (zie boven).

Resetten beeldbewerkingen

Wanneer je terug wilt naar de beginwaarden van het beeld (bv doorsnede- en grijswaarde verdeling waarin het beeld opkomt bij het openen van de vraag), druk je op de knop Reset (figuur 1).

Annotaties in beelden

In een beeld, volledige scan of 2D beeld, kan een structuur of afwijking zijn gemarkeerd, bv met een pijl, asterisk of doordat deze is ingekleurd. Je kan deze markering weghalen en weer laten verschijnen door op de letter “o” op het toetsenbord te klikken.

Vragen beantwoorden

Bij een beeldvraag is in de kolom links van de viewer de vraag te lezen. Er zijn meerdere vraagtypen mogelijk: meerkeuze-, meer uit meer-, aanwijs- en longmenuvragen. Hieronder volgt per vraagtype uitleg over hoe je deze moet beantwoorden.

Meerkeuzevragen

Bij een meerkeuzevraag zijn een aantal alternatieven als antwoord gegeven. Kies hieruit het beste antwoord door het vakje voor dit alternatief aan te vinken. Je kunt maar één antwoord geven.

Meer uit meer vragen

Bij een meer uit meer vraag is het mogelijk om meerdere alternatieven als antwoord te geven, dit in tegenstelling tot een meerkeuzevraag. Je kunt dus de vakjes voor meerdere alternatieven aanvinken.

Aanwijsvraag

Figuur 2 Aanwijsvraag

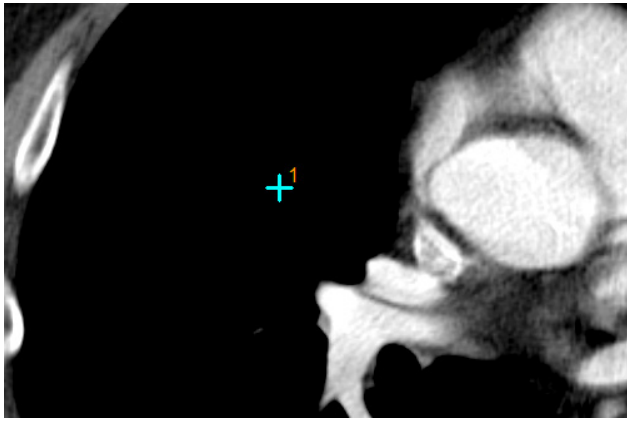


Bij een aanwijsvraag is het de bedoeling dat je de gevraagde structuur in een radiologisch beeld opzoekt en markeert met de marker (een plus).

Doe dit als volgt:

- Zoek eerst de structuur op. Bij een volledige CT-scans kun je op verschillende manieren bekijken (*scrollen, wisselen van richting en vensterinstelling*). Zie hiervoor de uitleg onder “Bekijken van volledige CT-scans” eerder in deze handleiding.
- Druk vervolgens op de knop “plaats marker” (figuur 2a). Met de linkermuisknop kun je nu de marker plaatsen.
- Klik in de afwijking of structuur die gevraagd wordt. Op deze plaats verschijnt nu een gekleurde plus (figuur 3). Door op een andere plaats in de scan te klikken, verplaats je ook de plus naar deze nieuwe plek. Als je tevreden bent over de plaats van de marker, **druk dan altijd eerst op “Bevestig locatie” (figuur 2b)!** Met de linkermuisknop kan je nu evt. weer scrollen door de CT-scan.
- De marker is nu in alle 3 de richtingen te zien (hij staat in de 3D ruimte).
- Als je de marker wilt verplaatsen, klik dan altijd eerst op de button “Plaats marker”
- Wanneer je tevreden bent over de plaats van de marker, **klik dan opnieuw op “bevestig locatie”**.
- Wanneer je aan het scrollen bent en je wilt terug naar de coupe waarin de plus staat, klik dan op “Ga naar coupe met markering”.

Figuur 3



Longmenuvraag

Bij een longmenuvraag is het de bedoeling dat je uit een lange lijst met alternatieven het juiste antwoord selecteert. Hiervoor is een dropdownmenu weergegeven. Soms zijn er twee dropdownmenu's weergegeven, bijvoorbeeld als gevraagd wordt naar twee mogelijke diagnoses of een differentiële diagnose.

Zoeken in een dropdownmenu:

- Type minimaal twee opeenvolgende letters van je antwoord in in het venster van het dropdownmenu en klik op de pijl om de lijst te openen. Alle termen die de ingetypte letters bevatten verschijnen nu in beeld.
- Kies een antwoord door op een alternatief uit de lijst te klikken.
- Het gekozen antwoord verschijnt nu in het dropdownmenu-venster en de lijst klapt weer dicht.
- Als je een ander antwoord wilt geven, haal je het antwoord in het venster weg en type je opnieuw enkele letters in.
- Zolang het woord of de lettercombinatie die in het venster staat niet in de lijst voorkomt, blijft het venster oranje.
- NB. Wanneer jouw antwoord er niet bij staat, dan kan je het ook niet als antwoord geven.

Colofon

Dit handboek maakt onderdeel uit van het project Toetsing in Beeld, gesubsidieerd door de Stichting SURF. Dit handboek is samengesteld door een samenwerking tussen de afdeling Radiologie, Image Science Institute, Anatomie en het Expertisecentrum van het UMC Utrecht en Departement Educatie, Faculteit Sociale Wetenschappen UU.

Voor de uitvoering van de onderzoeken en activiteiten is door vele personen een bijdrage geleverd, waarvoor onze dank. Onze speciale dank gaat uit naar: Pieter Jan van der Schoot (Functioneel beheerder bij I&I Onderwijscentrum UMC Utrecht); de onderwijsassistenten radiologie in het UMC Utrecht, in het bijzonder Iris ten Katen, Mohammed Attrach, Ufuk Kizilatis, Ralf Sprengers, Margiet de Haan, Beatrijs Seinstra, Saskia Jol; medewerkers van het KVO (Klinisch VaardigheidsOnderwijs), in het bijzonder Monique Helten, Janette van Barneveld, Margot Ligtenberg; Lisette van Bruggen (expertisecentrum, digitale toetsing); Bobby Stuijtzand (research master student, Departement Educatie, faculteit sociale wetenschappen UU); Onderwijssecretariaat Radiologie, Linda Ijskes en Amanda Groenewoud; Ignace Hooge (Psychologische Functie-leer); alle Utrechtse geneeskunde studenten en co-assistenten radiologie die hebben deelgenomen aan het onderzoek; de student-assistenten radiologie; de juniordocenten van de afdeling anatomie, die hebben ondersteund bij het snijzaalonderzoek; Marieke Krispijn (bedrijfsbureau divisie Beeld UMC Utrecht); Karin van Rijnbach en Roy Sanders (Multimedia divisie Beeld UMC Utrecht); alle medewerkers van Stichting Surf voor het mogelijk maken van het project Toetsing in Beeld, in het bijzonder Annette Peet, Lieke Rensink, Michiel van Geloven, Jenny de Werk, Anne Floor Erdman en Esther Arrindell.

Grafisch ontwerp, opmaak en drukwerk: Karin van Rijnbach, Multimedia divisie Beeld UMC Utrecht
Fotografie (figuur 5): Roy Sanders, Multimedia divisie Beeld UMC Utrecht
Grafisch ontwerp Omslag: Wilbert Bartels, ISI, divisie Beeld, UMC Utrecht

© 2014

1^e druk

Herziene versies zullen beschikbaar komen op website www.vquest.nl

Dit project is tot stand gekomen met steun van Stichting SURF, de organisatie die ICT vernieuwingen in het hoger onderwijs en onderzoek initieert, regisseert en stimuleert door onder meer het financieren van projecten. Meer informatie over SURF is te vinden op de website (www.surf.nl).