

Verbeteringen in digitaal toetsen UMC Utrecht

Een deel van de opleidingen die het UMC Utrecht verzorgt worden digitaal getoetst, waarbij het UMC Utrecht gebruik maakt van het toetsprogramma TestVision. Wat zijn de ervaringen en verbetermogelijkheden?

De keuze voor het toetsprogramma TestVision (TV), dat overigens ook wordt ingezet bij de specialistenopleidingen en opleidingen voor verpleegkundigen en operatieassistenten, is gedaan vanwege het gebruikersgemak en de duidelijke onderwijskundige structuur die in het programma is ingebouwd (de leerboom en toetsmatrijs). Een ander pluspunt is dat TV wordt geleverd door een Nederlands bedrijf waardoor de lijnen kort zijn.

In dit artikel staan de digitale toetsen bij blokken waarin studenten worden voorbereid op de coschappen centraal. Totaal zijn er zes coschapegebonden blokken, in vier daarvan wordt sinds 2003-2004 digitaal getoetst. In dit artikel wordt antwoord gegeven op de volgende vragen: Hoe waarderen studenten en docenten de huidige digitale toetsen? Wat zijn de wensen van docenten en afdelingen voor digitaal toetsen en het inrichten van een toetsbank? Zijn verbeteringen mogelijk en noodzakelijk? Wat voor ondersteuning hebben docenten nodig?

Kwaliteit

Binnen de opleiding geneeskunde wordt voor de summatieve toetsing van de blokken die voorbereiden op de coschappen gynaecologie, dermatologie, neurologie en oogheelkunde gebruik gemaakt van TV. Aanleiding voor het project was de wens van de afdeling OnderwijsTechnologie (OT) van de directie Onderwijs & Opleidingen (dOO) te onderzoeken of het huidige systeem voor digitale toetsing onbenutte mogelijkheden biedt en/of de kwaliteit van digitale toetsing kan worden geoptimaliseerd. Bijvoorbeeld door de toetsen beter te laten aansluiten bij de beroepspraktijk, zowel inhoudelijk, door de kwaliteit van de vragen te verbeteren, als digitaal, door optimaal gebruik te maken van de digitale mogelijkheden. Bij het ontwikkelen van TV is het opleidingskundig proces van het ontwikkelen, afnemen en analyseren van vragen en toetsen het uitgangspunt. Leerdoelen vormen de ruggengraat van TV.

In TV wordt hiervoor de term 'leereenheden' gebruikt. Anderen, zoals Draaijer et al (2004) gebruiken hiervoor de term 'onderwerp-/domeinindeling'. De leereenheden worden geordend in een zogenaamde leerboom, deze omvat het hele vakgebied. Alle vragen worden gekoppeld aan takken van de leerboom en voorzien van meta-informatie, bijvoorbeeld kennis/inzicht, moeilijkheidsgraad, formatief-summatief en vraagtype. Op basis van de leereenheden en meta-informatie worden specifieke toetsen gegenereerd, dit gebeurt via een toetsmatrijs. In de literatuur over toetsen wordt dit ook wel een toetsplan genoemd (zie Milius, 2007).

Gynaecologie

Het coschap gynaecologie duurt zes weken, het voorbereiden de blok wordt twaalf maal per jaar aangeboden en duurt vier weken. De meeste informatie in het blok is nieuw. In de toets worden vragen gesteld over de nieuw verworven inzichten en eerder behandelde onderdelen. In TV zitten driehonderd gynaecologievragen. Deze zijn niet gekoppeld aan een leerboom. Per toets worden er vijfendertig vragen geselecteerd die op basis van de kwalitatieve analyse uit TV 'opgeschoond' worden. De multiple choice (mc)-vragen zijn ontwikkeld door de coördinator van het coschap gynaecologie, in samenwerking met docenten. De reden voor het overgaan naar digitale toetsing was de toetsfrequentie en de correctietijd die bij digitale afname bespaard zou worden. Het invoeren van de vragen in TV en klaarzetten van de toetsen gebeurt door een medewerker van de afdeling OT.

De meeste toetsvragen zijn te typeren als korte casusbeschrijvingen met mc-vragen. De essentie van gynaecologie is 'het verhaal van de patiënt' dat als patiëntendossier in TV kan worden opgenomen (casus), compleet met eventueel beeldmateriaal (bijvoorbeeld echo), geluid (harttonen), labuitslagen, etc., waarbij verschillende vragen kunnen worden gesteld.

Lisette van Bruggen Magda Ritzen

Van Bruggen is onderwijskundig medewerker toetsing en beoordeling bij het Expertisecentrum voor Onderwijs en Opleiding van het UMC Utrecht. Ze is onder andere coördinator TestVision. Reacties naar: J.M.E.vanBruggen-4@umcutrecht.nl
Ritzen is adviseur hoger onderwijs en projectleider bij het IVLOS, Universiteit Utrecht. Ze werkt vooral aan projecten rondom onderwijs en ict. Reacties naar: m.ritzen@uu.nl

Dermatologie

Het coschap dermatologie duurt twee weken, het voorbereidende blok wordt vierentwintig keer per jaar aangeboden en duurt twee weken. De meeste informatie in het blok is nieuw, daarnaast komen alle eerder behandelde onderdelen aan bod. In TV zitten honderdtwintig nieuw ontwikkelde vragen (2009), daarvan zijn twee mc-toetsen gemaakt met elk zestig vragen die elkaar afwisselen. De nieuwe vragen zijn ontwikkeld door een AKO¹ na het volgen van een training over vraagconstructie². Ze zijn kwalitatief goed volgens de analyseregels van de psychometrie. Daarnaast zijn er honderdtwintig 'oude' vragen van wisselende kwaliteit die worden gebruikt bij herkansingen. Bij elke mc-vraag hoort een foto afkomstig van 'eigen patiënten'. Het invoeren van de oude vragen gebeurde door een medewerker van de afdeling OT, inmiddels heeft de AKO die taak overgenomen. De essentie van dermatologie is het statische visuele beeld: de overstap naar een digitaal toetsinstrument maakt het gebruik van kleurenfoto's vele malen goedkoper en de weergave kwalitatief beter.

Neurologie

Het coschap neurologie duurt zes weken, het voorbereidende blok wordt twaalf keer per jaar aangeboden en duurt twee weken. In het blok komen zowel nieuwe als reeds behandelde onderdelen aan bod. De mc-toets wordt afgenomen tijdens het coschap.

In TV zitten inmiddels ruim zeshonderd vragen die zijn gekoppeld aan een leerboom die is ingericht op basis van het 'Raamplan 2001 Artsopleiding'. Een deel van de vragen is veranderd, momenteel worden de vragen door de coördinator aangevuld, verbeterd en geactualiseerd. Het verbeteren gebeurt onder andere op basis van feedback van studenten en kwalitatieve analyses. Er wordt gestreefd naar een databank met vijftienhonderd goede toetsvragen. Het samenstellen van een toets, het klaarzetten voor afname en de statistische analyse gebeurt door medewerkers van de afdeling OT. De essentie van neurologie is 'het verhaal van de patiënt' in combinatie met neurologisch onderzoek. TV wordt nu vooral ingezet om de nakijktijd van docenten te beperken.

Oogheelkunde

Het coschap oogheelkunde duurt twee weken, het voorbereidende blok wordt vierentwintig keer per jaar aangeboden en duurt twee weken. In TV zitten honderdtwintig oogheelkunde-vragen die niet gekoppeld zijn aan een leerboom. De vragen zijn ontwikkeld door een arts-assistent die de training vraagconstructie heeft gevolgd. Er wordt gewerkt met drie verschillende vaste toetsen. Er is wel zicht op de kwaliteit van de vragen, maar daar zijn door tijdgebrek nog geen consequenties aan verbonden. Het voornemen is om nieuwe vragen te ontwikkelen en deze toe te voegen aan een toets, maar deze pas te

laten meetellen als ze volgens de kwalitatieve analyse goed scoren. De essentie van oogheelkunde is net als bij dermatologie het statische visuele beeld: de overstap naar een digitaal toetsinstrument maakt ook hier de toetsen goedkoper en het beeld kwalitatief beter.

Verbetermogelijkheden

Bovenstaande informatie is verzameld in gesprekken met de coördinatoren en enkele andere medewerkers van de verschillende betrokken disciplines. In deze gesprekken zijn ook verbeterwensen en -mogelijkheden geformuleerd:

- Uitbreiden van het bestand met toetsvragen.
- Meer variatie in vraagtypen, bijvoorbeeld matchingsvragen, hotspotvragen, meerdere antwoorden goed. Diversiteit in vraagtypen maken de toets levendiger en bieden docenten meer mogelijkheden parallelle vragen te maken.
- Kwaliteitsverbetering van vragen, zowel redactioneel³ als inhoudelijk.
- Vragen onderbrengen in een databasestructuur gekoppeld aan een leerboom en toetsmatris zodanig zodat automatische toetsgeneratie mogelijk is.

Bij gynaecologie en neurologie zouden de toetsen geoptimaliseerd en uitgebreid kunnen worden door te werken met beeldmateriaal en filmpjes. Bijvoorbeeld bij gynaecologie via een hartfilmpje van een baby in de buik tijdens de bevalling en bij neurologie via een filmpje van een lopende patiënt met een bepaalde stoornis.

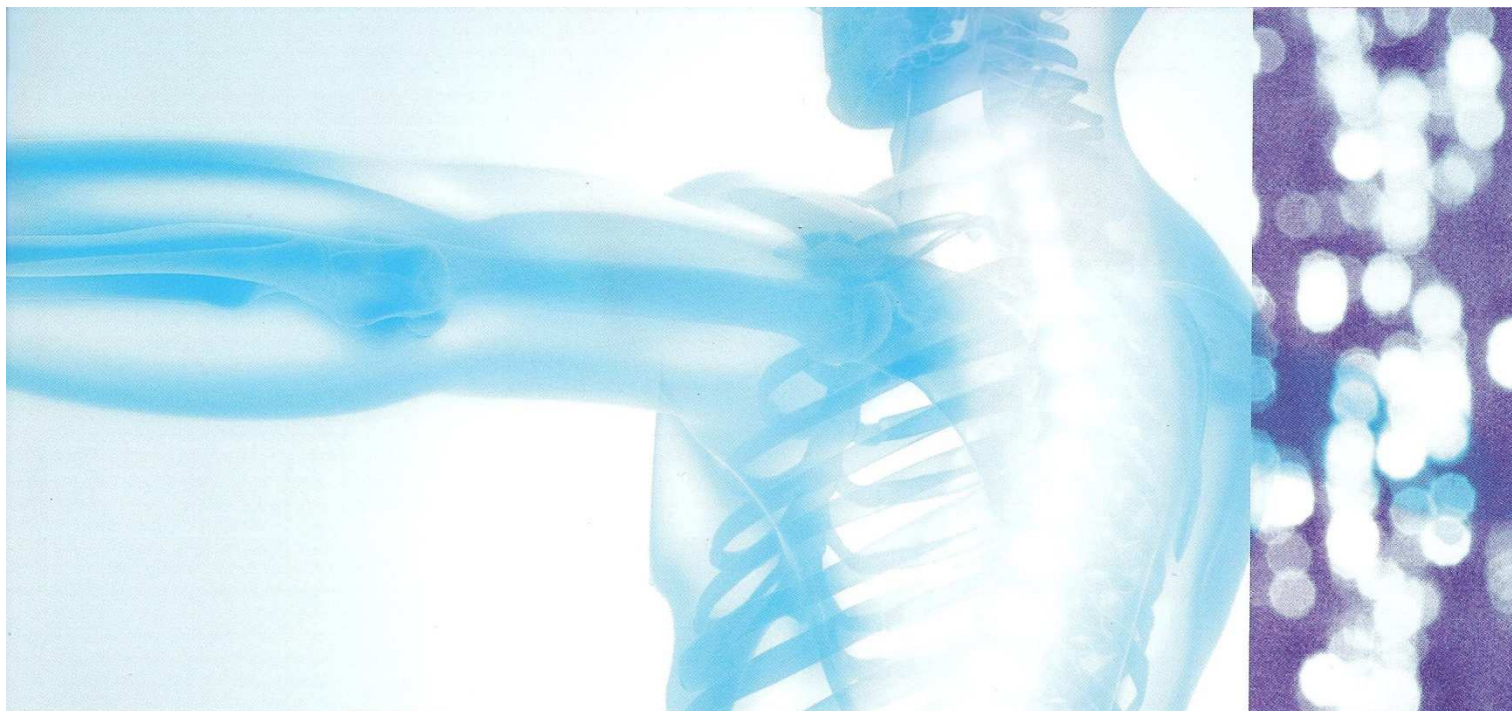
Daarnaast is door de afdeling OT een analyse van de toetsdatabase uitgevoerd. Hieruit kwam het volgende naar voren:

- Er worden alleen één-uit-meervragen en juist/onjuist-vragen gebruikt.
- Er worden meer kennis- dan inzichtvragen gesteld.
- Bij gynaecologie en neurologie wordt geen beeldmateriaal gebruikt.
- De leerboomstructuur en toetsmatris worden niet gebruikt zoals bedoeld in TV, daardoor wordt niet optimaal gebruik gemaakt van automatische toetsgeneratie en randomisering.
- Ongeveer een kwart van de vragen is toetstechnisch te verbeteren.

Studentenevaluaties

Uit recent onderzoek naar digitale toetsing blijkt dat studenten de digitale afname van een toets vanzelfsprekend vinden (Ravesloot et al 201X).

Om er achter te komen of studenten tevreden zijn met de digitale toetsen is in de periode oktober 2009 – maart 2010 bij alle coschapsgebonden toetsen een evaluatievragenlijst afgenomen. Uit de evaluatie bleek dat studenten [n= 277] tevreden zijn met de toets (score 6.7). Het enige verbeterpunt dat er enigszins uitspringt is het verhelderen van vragen door het toevoegen van



beeldmateriaal. Dit aspect komt ook terug in de verbeterwensen van de docenten van gynaecologie en neurologie.

Noch bij de student, noch bij de docent zijn er dringende redenen om de toetsen te verbeteren. Toch staan coördinatoren en docenten positief tegenover de voorgestelde verbeteringen en werken ze er graag aan mee. Omdat ze weinig tijd hebben, maken ze graag gebruik van de mogelijkheden die de afdeling OT biedt om de kwaliteit te verhogen: suggesties voor gebruik van andere vraagtypen, redactionele verbeteringen van toetsvragen, adviseren over en meewerken aan de inrichting van een leerboomstructuur en toetsmatrijs en het invoeren van benodigde gegevens en vragen in TV.

Suggesties voor alternatieve vraagtypen in relatie tot het toetsen van kennis of inzicht (Visser, 2009) en redactionele verbeteringen (zoals het aanpassen van onduidelijke of ontkennende vragen) kunnen door toetsdeskundigen worden ingevoerd die niet noodzakelijkerwijs kennis hebben van de inhoud.

Op dit moment zijn de leerboomstructuren vrij willekeurig ingericht en verschillend per discipline. Gynaecologie heeft geen leerboomstructuur, de leerbomen van dermatologie en neurologie zijn ingedeeld op basis van het 'Raamplan 2001 Artsenopleiding' en die van oogheelkunde op basis van een studieboek. Een toetsmatrijs gebaseerd op de leerboom wordt nauwelijks gebruikt. Een weloverwogen leerboomstructuur en toetsmatrijs leveren veel voordelen op (zie Dousma, Horsten et al 1997 en Hols-Elders et al 2008), zoals het automatisch genereren van verschillende toetsen die vergelijkbaar zijn wat betreft inhoud, niveau en kwaliteit.

Realiseren van verbeteringen

Over leerboomstructuur en toetsmatrijs zijn afspraken gemaakt met de coördinatoren van de betreffende disciplines. Mogelijke uitgangspunten voor een structuur zijn bediscussieerd en in relatie tot TV becommentarieerd door de medewerker van de afdeling OT. Voor dermatologie en gynaecologie heeft dit geleid tot een herziene leerboom en toetsmatrijs:

- Bij gynaecologie is een nieuwe leerboom opgezet voor het totale vakgebied en via de toetsmatrijs gekoppeld aan de leerdoelen van het coschapgebonden blok. Momenteel worden de huidige vragen in TV opnieuw beoordeeld en aan de leerboom gekoppeld. Ook worden oude vragen van papieren toetsen beoordeeld op bruikbaarheid. De nieuwe leerboom wordt gebruikt als basis voor verschillende toetsmatrijzen.
- Bij dermatologie wordt nu ook gewerkt met een leerboom, deze is eveneens ingedeeld op basis van het totale vakgebied en de onderwijsdoelen. De toetsmatrijs zal worden ingericht op basis van diezelfde leerboom.
- Bij oogheelkunde wordt gewerkt aan het inrichten van een leerboomstructuur die, zoals bij gynaecologie en dermatologie, gebaseerd is op het totale vakgebied en de onderwijsdoelen. De intentie is om de leerboom te gebruiken als basis voor het genereren van toetsen via toetsmatrijzen.
- Bij neurologie houdt men voorlopig vast aan leerboomstructuur op basis van het 'Raamplan 2001 Artsopleiding'. Dat geldt ook voor de toetsmatrijs, die is ingericht op basis van hetzelfde Raamplan in combinatie met het blokboek.

Bij de opleiding geneeskunde krijgen alle studenten een didactische training van twee weken. Daarnaast kunnen geïnteresseerde studenten kiezen voor een verdiepend keuzeblok op dit gebied: de onderwijsstage. Daarbij krijgen ze theoretische achtergronden, onder meer op het gebied van toetsing en/of vraagconstructie, verzorgen ze onderwijs aan eerstejaars studenten en voeren ze een onderwijsadviesproject uit voor een afdeling. Een afdeling (bijvoorbeeld neurologie of oogheelkunde) zou een dergelijke student kunnen vragen toetsvragen in TV te herzien of nieuwe vragen te ontwikkelen. Daarnaast zijn er afdelingen die een AKO hebben aangesteld met als hoofdtak onderwijs. Dermatologie is daarvan een voorbeeld. Dergelijke afdelingen gaan het meest voortvarend te werk met het reorganiseren en verbeteren van de toetsen.





Conclusie

De Utrechtse geneeskundedocenten en -studenten zijn over het algemeen tevreden met de digitale toetsen. Toch zijn er verbeterwensen, zoals uitbreiding van het aantal toetsvragen, hogere kwaliteit van de vragen en het gebruiken van een leerboomstructuur voor het genereren van toetsen. De verbeterwensen zijn vanwege tijdsdruk lastig te realiseren: docenten geven prioriteit aan het werk in de kliniek. Als ze tijd kunnen vrijmaken voor toetsing, besteden ze die in eerste instantie aan het ontwikkelen van nieuwe vragen.

Dankzij dit project staat het belang van een goede structuur in de leerboom en toetsmatrijs én het hebben van goede vragen beter op de kaart. Het inrichten van een leerboom via leerdelen of onderwerpen is een moeilijk proces. Gebleken is dat de afdeling OT hierin een stimulerende en faciliterende rol kan spelen. Met deze en andere technische ondersteuning en begeleiding van OT en parallel hieraan, het professionaliseren van docenten op het gebied van vraagconstructie, kan optimaal gebruik worden gemaakt van de digitale toetsomgeving TV en is er aandacht voor de kwaliteit van toetsing. OT kan geen bijdrage leveren aan de uitbreiding en inhoudelijke kwaliteit van de vragen. In dit project merkten we dat zowel studenten in hun onderwijsstage, als AKO's die geïnteresseerd zijn in onderwijs, steeds meer gaan doen in toetsontwikkeling op de afdelingen. Dat is goed nieuws, want zij zijn de nieuwe generatie docenten die vanuit de kliniek een bijdrage leveren aan onderwijs.

- 1 AKO: Arts Klinisch Onderwijs. Een afgestudeerde basisarts die naast zijn werk in de kliniek onderwijs geeft aan geneeskundestudenten.
- 2 Binnen het UMC Utrecht biedt het Expertisecentrum diverse docent-trainingen aan waaronder Toetsing I over vraagconstructie, Toetsing II over kwaliteitszorg en analyse van toetsen en Digitaal toetsen met TestVision.
- 3 Redactionele verbeteringen betreffen verbeteringen in de formulering van de vraag, bijvoorbeeld weghalen van ontkenningen, verwijderen van overbodige of vergezochte antwoordalternatieven, etc.



Literatuur

- Dousma T., Horsten A. et al. (1997). *Tentamineren*, Wolters-Noordhoff.
- Draaijer S. (2004). *Flexibilisering van Toetsing*. Utrecht, Digitale Universiteit.
- Hols-Elders W., Bloemendaal P., Bos N., Quaaq M., Sijstermans R., De Jong P. (2008) 'Twelve tips for computer-based assessment in medical education', *Medical Teacher*, 30:7, 673 – 678.
- Milius J. (2007). *Schriftelijk tentamineren. Een draaiboek voor docenten in het hoger onderwijs* (IVLOS-Reeks Hoger Onderwijs). Utrecht: IVLOS, Universiteit Utrecht.
- Ravesloot C.J., Cate Th., ten, Quaaq M., Ritzten M., Schaik, J.P.J. van (201X). *Computer-Based versus Paper-and-Pencil Testing: No Differences in Performance in a Randomized Study with a Radiology Test*. (paper in preparation).
- Visser D. *Het testen van kennis, inzicht en toepassing; welke toetsvragen zijn het meest geschikt?* In: Medisch-onderwijskundige verdieping, verslagen door onderwijsstagiairs 2008/2009. Uitgave Expertisecentrum voor Onderwijs en Opleiding, UMC Utrecht, november 2009.

