

## VAN DATA NAAR KLINISCHE IMPACT

# Het overbruggen van de AI-kloof in de radiologie



Laurens Topff

Ondanks indrukwekkende prestaties van AI-modellen in onderzoeksomgevingen blijft grootschalige klinische implementatie uit. De centrale vraag in dit promotieonderzoek luidt: hoe maken we de stap van experimentele modellen naar betrouwbare klinische hulpmiddelen?

Ondanks de aanzienlijke vooruitgang in AI-ontwikkelingen blijft er een opvallende kloof bestaan tussen onderzoeksresultaten en klinische implementatie. Hoewel talrijke AI-toepassingen hoge nauwkeurigheid hebben aangetoond in gecontroleerde omgevingen, bevindt hun adoptie in de dagelijkse radiologische praktijk zich nog in een vroeg stadium.

Om deze translationele kloof te overbruggen zijn enerzijds betrouwbare AI-modellen nodig die daadwerkelijk geschikt zijn voor klinisch gebruik, want veel modellen uit onderzoek falen bij externe validatie. Anderzijds moeten deze tools toegevoegde waarde bieden in de dagelijkse radiologische praktijk. Het eerste deel van dit proefschrift beschrijft de ontwikkeling van AI-toepassingen voor thoracale en oncologische beeldvorming. Hierin zijn methoden en uitdagingen bij het creëren en valideren van robuuste AI-modellen onderzocht. Het tweede deel gaat verder dan technische

neurale netwerkarchitecturen en trainingstechnieken om de prestaties van een model te verbeteren. Deze benadering staat bekend als 'model-centrische' AI. Er is echter een groeiend bewustzijn dat grote, diverse en goed geannoteerde datasets cruciaal zijn voor het succesvol vertalen van AI-modellen naar de klinische praktijk. In mijn onderzoek heb ik daarom een 'data-centrische' aanpak toegepast. Hierin verleg ik de primaire focus van modelverbetering naar het verbeteren van de kwaliteit van de gegevens. Het verschil tussen beide benaderingen is fundamenteel: waar de model-centrische strategie is gericht op het vinden van het best mogelijke model voor een gegeven dataset, richt de data-centrische benadering zich op het creëren van de hoogst mogelijke kwaliteit dataset voor modeltraining. Met deze benadering wil ik de generaliseerbaarheid en betrouwbaarheid van AI-modellen verbeteren.

Dit proefschrift beschouwt datakwaliteit als de hoeksteen van de ontwikkeling, evaluatie en implementatie van AI-mo-



## Opsporen hersenmetastasen

De ontwikkeling van een AI-model voor detectie en segmentatie van hersenmetastasen op MRI laat het belang van datakwaliteit goed zien. Eerder gepubliceerde modellen slaagden er niet in om kleine metastasen (< 6 mm) betrouwbaar te detecteren, vertonen een groot aantal foutpositieven, of generaliseren onvoldoende naar MRI-beelden van andere ziekenhuizen. Deze beperkingen ondermijnen de klinische bruikbaarheid van dergelijke modellen aanzienlijk.

Daarom hebben we vanuit de data-centrische benadering een generaliseerbaar AI-model ontwikkeld voor het detecteren en segmenteren van hersenmetastasen van alle formaten. Hiervoor hebben we een heterogene multicenter dataset van 1.985 scans van meer dan 30 verschillende MRI-scanners verzameld. Dit geeft

## 'De kwaliteit van de data bepaalt het klinische succes van AI'

beoordeling: het onderzoekt de toegevoegde waarde van AI-toepassingen in radiologische workflows.

### Datakwaliteit als fundament

Traditioneel wordt bij AI-ontwikkeling de prioriteit gelegd bij verbeteringen in

dellen in de radiologie. Het onderzoek maakt gebruik van nieuw verzamelde datasets van meerdere ziekenhuizen om de heterogeniteit te omvatten van beeldvormingsdata, met variaties in scanprotocollen, patiëntdemografie en klinische settings.

## De promotiedag



Op 2 april 2025 heb ik mijn proefschrift *Bridging the AI Gap in Radiology: From Data to Clinical Impact* succesvol verdedigd in de aula van de Universiteit Maastricht. Iedereen benadrukt altijd dat een promotie een bijzonder moment is. Toch realiseerde ik me pas op de dag zelf hoe betekenisvol dit is. Het gevoel van trots, opluchting en dankbaarheid tijdens de verdediging was groter dan ik had verwacht.

Ik wil mijn promotieteam hartelijk bedanken voor de begeleiding en het vertrouwen. Dank aan de oppositie voor de uitdagende vragen. Mijn paranimfen Kevin en Liliana, collega's, familie en vrienden bedank ik voor hun aanwezigheid en steun tijdens deze belangrijke mijlpaal. Ook wil ik mijn collega Petra de Koekoek bedanken, die met haar camera deze onvergetelijke dag perfect heeft vastgelegd.



een goede weergave van de variabiliteit van sequenties uit de klinische praktijk. De gegevens zijn op systematische wijze gecureerd en geannoteerd. Voor het maken van de referentiestandaard hebben radiologen de metastasen zorgvuldig gesegmenteerd, met uitgebreide kwaliteitscontroles.

Het model toonde een hoge sensitiviteit van 97,4% voor het detecteren van hersenmetastasen op externe data, inclusief voor de kleinste laesies van minder dan 3 mm in diameter (93,3%). Het aantal foutpositieven was acceptabel met een gemiddeld aantal van 0,6 per patiënt, waarbij de meerderheid van scans geen foutpositieven vertoonde. De segmentatieprestatie was erg goed met een *Dice Similarity Coefficient* van 0,90.

### Workflowintegratie

Naast modelontwikkeling is ook de toepassing in de praktijk van belang. Het tweede deel van het proefschrift richt zich op de integratie van AI-toepassingen in radiologische werkprocessen. Hierbij is gekeken naar de toegevoegde waarde van commercieel beschikbare AI-tools. Het herkennen van optimale interventiepunten binnen de radiologische workflow waar AI waarde toevoegt en verstoringen minimaliseert, is belangrijk.

Om gemiste bevindingen in het radiologische verslag op te sporen, hebben we een evaluatie gedaan met dubbele lezing van thoraxfoto's met behulp van AI. Het systeem heeft zowel de beelden als de bijbehorende verslagen geanalyseerd om discrepanties te herkennen die kunnen wijzen op gemiste bevindingen. In dit onderzoek van 25.104 thoraxfoto's uit 2 Nederlandse ziekenhuizen zijn klinisch relevante gemiste bevindingen gevonden in slechts 0,1% van de gevallen. Deze bestonden uit niet-gerapporteerde longnoduli, pneumothoraxen en consolidaties.

### Real-world evaluatie

Real-world evaluaties zijn essentieel voor het beoordelen van de klinische impact van AI-tools, aangezien zij inzicht verschaffen in de wijze waarop deze toepassingen presteren onder daadwerkelijke praktijkomstandigheden en interageren met bestaande werkprocessen. Een AI-toepassing voor het detecteren en prioriteren van incidentele longembolie op routine-CT-scans is prospectief onderzocht in een praktijk met een achterstand in niet-verslagen onderzoeken. De software had een hoge diagnostische ►

nauwkeurigheid met een sensitiviteit van 91,6% en specificiteit van 99,7% voor het detecteren van incidentele longembolie. Het klinische voordeel was de significante verkorting van de tijd tot diagnose van incidentele longembolie van meerdere dagen naar 1 uur.

essentieel om een naadloze integratie van AI-resultaten in de praktijk mogelijk te maken.

### Conclusie

Dit proefschrift heeft bijgedragen aan het overbruggen van de translationele

dige integratie in de kliniek én oog voor gebruiksgemak kan AI haar volledige potentieel in de radiologie waarmaken.

## ‘Real-world evaluaties zijn de lakmoesproef voor AI-modellen’

### Toekomstperspectief

Een focus op datakwaliteit blijft van belang voor het succesvol ontwikkelen van de volgende generatie, wellicht multimodale, AI-toepassingen in de radiologie. Het maximaliseren van de waarde van beperkt beschikbare medische data door goed ontworpen, efficiënte datapijplijnen, zal de ontwikkeling van betrouwbare AI-modellen voor klinisch gebruik bevorderen. Om optimaal gebruik te kunnen maken van AI-mogelijkheden zullen bestaande radiologische workflows herzien moeten worden en is de implementatie van interoperabiliteitsstandaarden

kloof van AI in de radiologie door zowel het potentieel als de uitdagingen van modelontwikkeling en implementatie te onderzoeken. Systematische datacuratie, rigoureuze externe validatie en multidisciplinaire samenwerking zijn essentieel voor het vertalen van AI-onderzoek naar de klinische praktijk. Alleen met zorgvul-



Het digitale proefschrift is beschikbaar op <https://doi.org/10.26481/dis.20250402lt> of via het scannen van deze QR-code.

Amsterdam, 2 juli 2025

### dr. Laurens Topff

radioloog en Medical Information Officer, Antoni van Leeuwenhoek

*Met veel dank aan mijn promotor:*  
prof. dr. Regina G.H. Beets-Tan,  
Universiteit Maastricht/Antoni van Leeuwenhoek

*Mijn copromotoren:*  
prof. dr. Erik R. Ranschaert,  
Universiteit Gent/St. Nikolaus Hospital  
dr. Jacob J. Visser,  
Erasmus MC