

Mri-Based Radiomics in Breast Cancer

Citation for published version (APA):

Granzier, R. W. Y. (2022). *Mri-Based Radiomics in Breast Cancer: Optimization and Prediction*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20220922rg>

Document status and date:

Published: 01/01/2022

DOI:

[10.26481/dis.20220922rg](https://doi.org/10.26481/dis.20220922rg)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

CHAPTER 9

Summary

This thesis consists of two parts that have the common goal of advancing personalized breast cancer care through MRI-based radiomics. Part one of this thesis investigated the use of MRI-based radiomics for prediction purposes in the treatment of breast cancer patients. Part two of this thesis focused on optimizations of MRI-based radiomics.

Part I – MRI-based radiomics for prediction purposes in breast cancer patients

In **chapter 3**, a descriptive systematic review was performed to create an overview of studies investigating the value of MRI-based radiomics for predicting tumor response to neoadjuvant systemic therapy in breast cancer patients. A total of 16 studies were included, examining data from 1,736 in total patients. The methodological quality of the included articles was judged as poor with an average radiomics quality score (RQS) of just 11% (range 0 – 41.2%). This was mainly due to the lack of external validation. In addition, radiomics methodologies showed large differences between studies, especially for tumor segmentation, feature selection, and model development, resulting in heterogeneous results that could not be compared. Nevertheless, the majority of the included articles showed promising results. Looking at the individual features, *entropy* emerged as the best performing feature with AUC values ranging from 0.83 to 0.85. The best performing multivariate prediction model scored a validation AUC value of 0.94. Based on these results, it was concluded that there is a need for standardization of the radiomics methodology to obtain comparable results in order to make further progress in this research area.

Chapter 4 investigated the possibility to predict pathological complete tumor response to neoadjuvant systemic therapy based on pretreatment MRI exams. A total of 292 breast cancer patients, with 320 breast tumors, were included in the analysis. Since the data was collected in two hospitals with five different MRI scanners and varying scanning protocols, three different strategies were used to split the data into training and validation cohorts. Radiomics, clinical, and combined models were developed and validated. The results showed that the radiomics models had no added value in predicting pathologic complete tumor response to neoadjuvant systemic therapy compared to the clinical models, nor did the combined models significantly outperform the clinical models. However, it should be noted that the effect of using data from different hospitals (with different MRI scanners and different scanning protocols on the extracted radiomic features), is still unknown. It was therefore concluded that these effects should first be investigated to determine whether further research on MRI-based radiomics for the prediction of pathologic complete response in breast cancer patients is useful.

Chapter 5 investigated whether radiomics analysis of T2-weighted dedicated axillary MRI exams can contribute to the improved diagnostic accuracy of the MRI for the prediction of axillary lymph node metastases. In this study, 511 axillary lymph nodes from 75 breast cancer patients were examined. Before the start of the radiomics analysis, all axillary lymph nodes were manually segmented in three dimensions and matched with pathology, after which 105 original radiomics features were extracted per lymph node. To validate the results, the data cohort was split into training and validation cohorts, this cohort split was performed 100 times. Each cohort split resulted in a different selection of radiomic features and with that in different AUC values. The performance of the clinical and radiomics models showed a wide range of AUC values of 0.41 – 0.74 and 0.48 – 0.89 in the training cohorts, respectively, and 0.30 – 0.98 and 0.37 – 0.99 in the validation cohorts, respectively. Based on these results, it was not possible to develop a definitive prediction model. It was concluded that radiomics analysis of dedicated T2-weighted axillary MRI exams did not contribute to the prediction of axillary lymph node metastases in breast cancer patients.

Part II – Optimization in MRI-based radiomics

Chapter 6 determined the robustness of radiomic features extracted using two commonly used radiomics software packages (RadiomiX and Pyradiomics) with respect to variability in manual breast tumor segmentations on MRI exams. A total of 129 breast tumors were segmented manually in three dimensions, by four observers: a dedicated breast radiologist, a resident, a Ph.D. candidate, and a medical student. The segmentation variability was measured using the volumetric Dice Similarity Coefficient with a mean value of 0.81 (range 0.19 – 0.96) indicating a good overlap of the breast tumor segmentations. The robustness of features was measured by the intraclass correlation coefficient. In total, 41.6% and 32.8% of all RadiomiX and Pyradiomics features, respectively, were identified as robust, independent of inter-observer manual segmentation variability.

In **chapter 7**, an MRI test-retest study was performed to assess the repeatability of radiomic features extracted from breast MRI exams. In total, eleven healthy volunteers were scanned on the same 1.5 Tesla MRI scanner in the MUMC+ using an identical scan protocol consisting of T1-weighted images, T2-weighted images, and diffusion-weighted images with corresponding ADC maps. For each healthy volunteer, 18 MRI exams were scanned on two separate days using three different test-retest strategies. From each MRI exam, the right breast was manually segmented in three dimensions where after 91 original radiomic features were extracted. Feature repeatability has been determined for features extracted from the original unprocessed images and pre-processed images. For the T1-

weighted images, the original unprocessed images were found to have the highest percentage of repeatable features, at 16.5%. The T2-weighted images showed the highest percentage of repeatable features after image pre-processing by z-score normalization, at 28.6%. The ADC maps, like the T1-weighted images, showed the highest percentage of repeatable features on the unprocessed original images, with only 8.8% of the features being repeatable. These results showed that the percentage of repeatable features in this specific setting is limited and varied per MRI sequences and image pre-processing procedure.

Samenvatting

Dit proefschrift bestaat uit twee delen met als gemeenschappelijk doel een stap verder te zetten in gepersonaliseerde borstkankerzorg met behulp van op MRI gebaseerde radiomics. Deel 1 van dit proefschrift onderzoekt het gebruik van op MRI gebaseerde radiomics voor de ontwikkeling van predictiemodellen. In deel 2 van dit proefschrift werd gekeken naar de optimalisatie van MRI gebaseerde radiomics.

Deel 1 – MRI gebaseerde radiomics voor de ontwikkeling van predictiemodellen

In **hoofdstuk 3** werd een beschrijvende systematische review uitgevoerd naar de voorspellende waarde van MRI-gebaseerde radiomics voor tumor respons op neo-adjuvante systemische therapie bij borstkanker patiënten. In totaal werden 16 studies geïnccludeerd waarin data van in totaal 1,736 patiënten werden onderzocht. Met een gemiddelde radiomics quality score (RQS) van slechts 11% (range 0 - 41.2%) werd de methodologische kwaliteit van de geïnccludeerde artikelen als slecht beoordeeld. Dit werd met name veroorzaakt door het ontbreken van externe validatie. Er waren grote verschillen te zien in gebruikte radiomics methodologieën tussen de studies, met name in de tumor segmentatie, feature selectie en model ontwikkeling; dit wat resulteerde in heterogene, niet vergelijkbare resultaten. Desalniettemin, liet de meerderheid van de geïnccludeerde artikelen veel belovende resultaten zien. Het individuele feature *entropy* liet de hoogste AUC waarde zien met een variatie van 0.83 tot 0.85. Het best presterende predictiemodel, gebruikmakend van meerdere features, had een validatie AUC waarde van 0.94. Op basis van de verkregen resultaten werd geconcludeerd dat er een noodzaak is om de radiomics methodologie verder te standaardiseren om beter vergelijkbare resultaten te verkrijgen en zo vooruitgang te kunnen boeken op dit onderzoeksgebied.

In **hoofdstuk 4** werd onderzocht of het mogelijk is om door middel van radiomics te voorspellen of een tumor in de borst volledig zal verdwijnen (pathologische complete tumorrespons) na het geven van neo-adjuvante systemische therapie, op basis van MRI scans die voor aanvang van de behandeling gemaakt zijn. In dit onderzoek werden MRI scans geanalyseerd van 290 patiënten met in totaal 320 borsttumoren. Omdat de data werden verzameld uit twee ziekenhuizen waarbij er werd gescand met verschillende MRI scanners en scan protocollen, werden de te analyseren data opgesplitst in trainings- en validatie cohorten met behulp van drie verschillende strategieën. Radiomics, klinische en gecombineerde modellen werden ontwikkeld en gevalideerd. De radiomics modellen hadden geen

toegevoegde waarde in het voorspellen van pathologisch complete tumorrespons op neo-adjuvante systemische therapie op de klinische modellen. Ook presteerden de gecombineerde modellen niet significant beter dan de klinische modellen. Het effect van het gebruik van data van verschillende ziekenhuizen, verschillende MRI scanners en verschillende scanprotocollen op de geëxtraheerde radiomics features is nog steeds onbekend. Daarom werd geconcludeerd dat dit effect eerst onderzocht moest worden om te bepalen of verder onderzoek naar op MRI gebaseerde radiomics voor de voorspelling van pathologische complete respons bij borstkankerpatiënten zinvol is.

In **hoofdstuk 5** werd onderzocht of radiomics analyse van de T2-gewogen beelden van de axillaire MRI kan bijdragen aan een verbeterde diagnostische accuratesse van de MRI voor de predictie van axillaire lymfekliermetastasen. In deze studie werden 511 axillaire lymfeklieren van 75 borstkanker patiënten onderzocht. Voor de start van de radiomics analysis werden alle axillaire lymfeklieren handmatig in drie dimensies gesegmenteerd en gematcht met de pathologie, waarna per lymfeklier 105 originele radiomics features werden geëxtraheerd. Om gevonden resultaten te valideren werd het data cohort opgesplitst in een training en validatie cohort, deze cohort-splitsing werd 100 keer uitgevoerd. Elke cohort-split resulteerde in een verschillende selectie van radiomics features en daarmee ook in verschillende AUC waarden. Om de radiomics resultaten te vergelijken werden er ook klinische modellen geanalyseerd. De klinische en radiomics modellen resulteerden in een grote variatie van AUC waarden van respectievelijk 0.41 – 0.74 en 0.48-0.89 in de trainingscohorten en respectievelijk 0.30 – 0.98 en 0.37 – 0.99 in de validatie cohorten. Op basis van deze resultaten was het niet mogelijk om een definitief predictiemodel te ontwikkelen. De conclusie was dan ook dat radiomics analyse van een axillaire MRI, op basis van de gevonden resultaten, geen bijdrage levert aan de voorspelling van axillaire lymfekliermetastasen in borstkankerpatiënten.

Deel 2 – Optimalisatie van MRI gebaseerde radiomics

In **Hoofdstuk 6** werd de robuustheid van radiomics features via twee veelgebruikte radiomics software pakketten (RadiomiX en Pyradiomics) beoordeeld met betrekking tot variabiliteit in handmatig gesegmenteerde borsttumoren op MRI scans. In totaal werden er 129 borsttumoren handmatig, in drie dimensies, gesegmenteerd door vier personen: een radioloog gespecialiseerd in borsttumoren, een radioloog in opleiding, een promovendus en een geneeskundestudent. De segmentatie variabiliteit werd gemeten aan de hand van de volumetrische *Dice Similarity Coefficient*. De gemiddelde waarde hiervan was 0.81 (range 0.19 - 0.96) wat

een goede overlap van de borsttumor segmentaties indiceerde. De robuustheid van features werd gemeten aan de hand van de intraclass correlatie coëfficiënt. In totaal werden respectievelijk 41.6% en 32.8% van alle RadiomiX en Pyradiomics features als robuust geïdentificeerd. Deze zijn derhalve onafhankelijk van variabiliteit verkregen door handmatig uitgevoerde borsttumor segmentaties.

In **hoofdstuk 7** werd een MRI test-retest studie uitgevoerd voor de beoordeling van de herhaalbaarheid van radiomics features geëxtraheerd van MRI scans van de borsten. In totaal werden 11 gezonde vrijwilligers gescand op dezelfde 1.5 Tesla MRI scanner in het MUMC+ gebruikmakend van een identiek scanprotocol bestaande uit T1-gewogen beelden, T2-gewogen beelden, en diffusie-gewogen beelden met bijbehorende ADC-map. Van elke gezonde vrijwilliger werden 18 MRI scans gemaakt middels drie verschillende test-retest settings op twee verschillende dagen. Van iedere scan werd de rechter borst in drie dimensies gesegmenteerd, waarna per borst 91 originele radiomics features werden geëxtraheerd uit de MRI. De feature herhaalbaarheid werd getest voor zowel features geëxtraheerd uit de originele onbewerkte beelden als uit de voorbewerkte beelden. Voor de T1-gewogen beelden bleken de onbewerkte beelden het hoogste percentage herhaalbare features te hebben, namelijk 16.5%. Daarentegen vertoonden de T2-gewogen beelden het hoogste percentage herhaalbare features nadat de beelden waren voorbewerkt middels z-score normalisatie, namelijk 28.6%. De ADC-maps vertoonden, net als de T1-gewogen beelden, het hoogste percentage herhaalbare feature op de onbewerkte beelden, met slechts 8.8% van de features die herhaalbaar waren. Deze resultaten toonden aan dat het percentage herhaalbare features in deze specifieke setting beperkt is en ook varieert per MRI-sequentie en per pre-processing procedure.