

Quantitative methods for improved error detection in dose-guided radiotherapy

Citation for published version (APA):

Wolfs, C. J. A. (2020). *Quantitative methods for improved error detection in dose-guided radiotherapy*. Ipskamp Printing BV. <https://doi.org/10.26481/dis.20200925cw>

Document status and date:

Published: 01/01/2020

DOI:

[10.26481/dis.20200925cw](https://doi.org/10.26481/dis.20200925cw)

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.

Summary

Radiotherapy is one of the main treatment options for many cancer types. Over the years, radiotherapy treatment delivery has become more precise and complex, with the aim of delivering a high radiation dose to the tumor and as little dose as possible to surrounding healthy tissues. With the increasing complexity of radiotherapy, it has become increasingly important to verify that the desired radiation dose is delivered as planned. This thesis focuses on dose-guided radiotherapy (DGRT) using portal dosimetry with the electronic portal imaging device (EPID) as a method for treatment verification. Measured dose distributions with the EPID can be compared to planned dose distributions, such that relevant deviations in the treatment can be detected. The aim is to flag treatments in which relevant dosimetric errors occur, such that the treatment can be adapted. However, automatic error detection and consequently patient selection for adaptive radiotherapy using portal dosimetry remains difficult. One reason for this is the unknown uncertainty of the portal dosimetry methods, while another reason is the lack of objective triggers for adaptation and the use of simple decision protocols (often consisting of an action threshold on a single metric extracted from dose comparisons) that cannot take the multidimensionality of the EPID dose distributions into account.

The overall aim of this thesis was to investigate and improve the error detection capabilities of portal dosimetry methods. To this end, the thesis is divided into two parts. The first part concerns uncertainties and sensitivity of portal dosimetry methods. In **Chapter 2**, a framework for analyzing uncertainties of portal dosimetry methods was established and this framework was applied to a larger dataset in **Chapter 3**, showing that the uncertainty is influenced by the presence of small radiation fields, which are becoming more common as treatment complexity increases. Therefore, it is important to incorporate this uncertainty in decision protocols for adaptive radiotherapy. The sensitivity and specificity of several portal dosimetry methods for detecting simulated anatomical changes using a simple threshold based classification method was investigated in **Chapter 4**, where it was shown that different combinations of portal dosimetry methods, comparison criteria and action thresholds are optimal for detecting different types of errors. This differs from current clinical practice, where the same comparison criteria and action threshold are used for all cases.

The second part of this thesis focuses on advanced methods for error detection and classification, applying artificial intelligence (AI) to portal dosimetry. AI techniques can take more information into account than the currently used action threshold classification methods. Several AI methods and classification problems were investigated to show the potential of AI-based error detection for portal dosimetry. **Chapter 5** provides a proof of concept of a deep learning model for detecting relevant dose errors in 2D dose comparisons based on portal dosimetry simulations. This shows that it is possible to relate 2D information to differences in metrics based on the 3D dose in the target structure by using a convolutional neural network (CNN). This relationship between 2D and 3D dose information is notoriously difficult to find

with traditional action threshold based classification methods. In **Chapter 6**, this work was extended to portal dosimetry measurements obtained during patient treatment, and includes dose metrics for organs-at-risk (OARs) as well. Here it was also possible to relate 2D dose comparisons to metrics extracted from 3D dose distributions, although generalizability of the model to new data is an issue that needs to be addressed in future work. **Chapter 7** took a different approach, as the aim was to classify the type of error that occurred. This simulation study showed that deep learning is a promising powerful tool for detecting types of treatment errors with portal dosimetry. Errors can be detected to a high level of detail, and this approach can provide additional information not currently available from portal dosimetry in clinical practice. In **Chapter 8**, a machine learning model for classifying severity of anatomical changes developed at CHU de Québec (Canada) was validated using portal dosimetry data from Maastricht. This study showed that it is difficult to use such a model on clinical data from a different institute due to differences in data acquisition and clinical interpretation. There was, however, a relationship between model classification (which was based on metrics extracted from 2D dose comparisons) and differences in 3D dose metrics for certain OARs, which can be useful for selecting patients that may need dose recalculation or adaptation of their treatment plan.

Overall, portal dosimetry for DGRT can greatly benefit from improved error detection methods. This thesis has contributed to this improvement, by providing a framework for uncertainty analysis of portal dosimetry methods, by examining the sensitivity and specificity for action threshold based classification methods for various portal dosimetry methods, and by investigating various advanced AI algorithms and error classification problems.

Samenvatting

Radiotherapie is een belangrijke behandeloptie voor veel kankertypes. In de loop der jaren is de toediening van radiotherapiebehandelingen nauwkeuriger en complexer geworden, met als doel het toedienen van een hoge stralingsdosis aan de tumor en zo min mogelijk aan omliggende gezonde weefsels. Met de toenemende complexiteit van radiotherapie, wordt het steeds belangrijker om te verifiëren dat de gewenste stralingsdosis wordt afgeleverd zoals gepland. Dit proefschrift richt zich op dosis-gestuurde radiotherapie (DGRT) met behulp van dosimetrie met elektronische röntgencamera's (electronic portal imaging device: EPID) als methode voor behandelverificatie. Gemeten dosisverdelingen met de EPID kunnen worden vergeleken met geplande dosisverdelingen, zodat relevante afwijkingen in de behandeling kunnen worden opgespoord. Het doel is behandelingen te identificeren waarbij relevante dosimetrische fouten optreden, zodat de behandeling kan worden aangepast. Automatische foutdetectie met deze verificatie methode, en daardoor patiëntselectie voor adaptieve radiotherapie, blijft een probleem. Eén reden hiervoor is de onbekende onzekerheid van de EPID dosimetrie methoden, terwijl een andere reden het gebrek aan objectieve grenswaarden voor aanpassing van de behandeling is en daarbij het gebruik van eenvoudige beslisprotocollen (vaak bestaand uit een drempelwaarde op een enkele variabele, afgeleid van dosisvergelijkingen) die geen rekening kunnen houden met de multi-dimensionaliteit van de EPID-dosisverdelingen.

Het doel van dit proefschrift was het onderzoeken en verbeteren van foutdetectie met EPID dosimetrie methoden. Daartoe is het proefschrift verdeeld in twee delen. Het eerste deel betreft de onzekerheden en sensitiviteit van EPID dosimetrie methoden. In **Hoofdstuk 2** werd een kader voor het analyseren van onzekerheden van EPID dosimetrie methoden opgesteld en dit kader werd toegepast op een grotere dataset in **Hoofdstuk 3**, waarin werd aangetoond dat de onzekerheid wordt beïnvloed door de aanwezigheid van kleine stralingsvelden, die steeds vaker voorkomen naarmate de complexiteit van de behandeling toeneemt. Het is daarom belangrijk om deze onzekerheid op te nemen in beslisprotocollen voor adaptieve radiotherapie. De sensitiviteit en specificiteit van verschillende EPID dosimetrie methoden voor het opsporen van gesimuleerde anatomische veranderingen met behulp van een eenvoudige classificatiemethode gebaseerd op drempelwaarden werd onderzocht in **Hoofdstuk 4**. Daar werd aangetoond dat verschillende combinaties van EPID dosimetrie methoden, vergelijkingscriteria en drempelwaarden optimaal zijn voor het detecteren van verschillende typen fouten. Dit verschilt van de huidige klinische praktijk, waar voor alle gevallen dezelfde vergelijkingscriteria en drempelwaarde worden gebruikt.

Het tweede deel van het proefschrift richt zich op geavanceerde methoden voor foutdetectie en classificatie, waarbij kunstmatige intelligentie (KI) wordt toegepast op EPID dosimetrie. KI technieken kunnen meer informatie verwerken dan de momenteel gebruikte classificatiemethoden gebaseerd op drempelwaarden. Verschillende KI methoden en classificatieproblemen werden onderzocht om de potentie van op KI

gebaseerde foutdetectie voor EPID dosimetrie aan te tonen. **Hoofdstuk 5** biedt een 'proof of concept' van een deep learning model voor het detecteren van relevante dosisfouten in 2D dosisvergelijkingen op basis van EPID dosimetrie simulaties. Dit laat zien dat het mogelijk is om 2D informatie te relateren aan verschillen in variabelen afgeleid van de 3D dosis in de doelstructuur/tumor met behulp van een convolutioneel neurale netwerk (CNN). Deze relatie tussen 2D en 3D dosisinformatie is moeilijk te vinden met traditionele classificatiemethoden gebaseerd op drempelwaarden. In **Hoofdstuk 6** werd dit werk uitgebreid met EPID dosimetrie metingen die werden verkregen tijdens de behandeling van patiënten, en zijn variabelen voor risicoorganen (organs-at-risk: OARs) ook meegenomen. Hier was het ook mogelijk om 2D dosisvergelijkingen te relateren aan variabelen afgeleid van 3D dosisverdelingen, hoewel de generaliseerbaarheid van het model naar nieuwe data een probleem is dat in de toekomst verder moet worden onderzocht. **Hoofdstuk 7** gebruikte een andere aanpak, waarbij het doel was om het type fout dat zich voordeed te classificeren. Deze simulatiestudie toonde aan dat deep learning een veelbelovend krachtig hulpmiddel is voor het detecteren van verschillende soorten behandelingsfouten met EPID dosimetrie. Fouten kunnen tot in detail worden opgespoord en deze werkwijze kan aanvullende informatie opleveren die momenteel in de klinische praktijk niet beschikbaar is vanuit EPID dosimetrie. In **Hoofdstuk 8** werd een machine learning model voor het classificeren van de ernst van anatomische veranderingen, ontwikkeld aan CHU de Québec (Canada), gevalideerd met behulp van EPID dosimetrie metingen van Maastricht. Deze studie toonde aan dat het moeilijk is om dit model te gebruiken op klinische data van een ander instituut, te wijten aan verschillen in data acquisitie en klinische interpretatie. Er was echter een verband tussen modelclassificatie (die was gebaseerd op variabelen afgeleid van 2D dosisvergelijkingen) en verschillen in variabelen afgeleid van de 3D dosis voor bepaalde OARs, wat nuttig kan zijn voor selectie van patiënten die mogelijk een dosisherberekening of aanpassing van hun behandelplan nodig hebben.

In het algemeen kan EPID dosimetrie voor DGRT aanzienlijk voordeel halen uit verbeterde methoden voor foutdetectie. Dit proefschrift heeft bijgedragen aan deze verbetering door een kader te bieden voor analyse van onzekerheden van EPID dosimetrie methoden, door de sensitiviteit en specificiteit te onderzoeken van classificatiemethoden gebaseerd op drempelwaarden voor verschillende EPID dosimetrie methoden, en door verschillende geavanceerde KI algoritmen en foutclassificatie problemen te onderzoeken.