



SUMMARY /
SAMENVATTING

Summary

Radiation oncology is in trouble. In fact, healthcare in general is. The spending on healthcare is rapidly rising, leading to a non-sustainable situation [1]. An aging population can partly explain the increasing costs, but overtreatment plays a large role as well [2]. This particularly happens due to defensive healthcare [3]. On the other hand, overtreatment also occurs due to the longevity of clinical protocols, which can cause inefficiency, as does the technology-push from healthcare industry.

One of the options to reduce the costs is to apply cost-effectiveness analyses or Healthcare Technology Assessments to inform policy and decision makers about the rightful use of (new) technology.

To this end, the development of personalized decision support systems was investigated, including cost-effectiveness in the field of proton therapy. The following hypotheses were addressed and affirmed:

- ◆ Rapid Learning Healthcare (RLHC) can improve clinical research and decision-making.
- ◆ Advances in Healthcare Information Technology (HIT) enable international data sharing while preserving patient privacy.
- ◆ Combining *in silico* planning comparison studies with multifactorial prediction models enables cost-effective application of limited, expensive resources, as introduced for proton therapy.

The basic concept of this thesis is the use of RLHC to complement evidence-based medicine with prediction modelling to provide clinical decision support for individualised cancer care (**Chapter 2**). RLHC applies advanced HIT to aggregate disparate clinical information sources into a data warehouse that presents a complete, integrated disease-oriented view of medical data for research purposes. This chapter presents a general overview of some of the techniques that were used in the following work.

Chapter 3 demonstrates how institutions can benefit from this improved presentation of patient information to improve efficiency and quality of data collection for clinical trials. Furthermore, RLHC uses machine learning principles to mine the data for clinically relevant relations to build data-driven complication prediction models.

Large amounts of data are needed for the modelling and validation of these models, which is unavailable from single institutions. Furthermore, to increase

information heterogeneity, which improves the predictive value of models, and to allow for external validation it is imperative that foreign data is available.

Chapter 4 presents the necessary data-sharing techniques that unlock the medical data silos in a privacy-preserving manner by installing ‘connectors’ inside the hospital’s network firewalls. Furthermore, by transforming the data into internationally standardised semantic interoperable data models, computer systems are able to translate data into knowledge. Furthermore, medical records in local languages are converted into international, unique codes. This means that the data can be reused in foreign countries with equal meaning and value, which allows for federated or distributed learning of complication prediction models [4].

In advance of a federated network, which is very advanced but not common practice yet, the framework of MISTIR (mistir.info) uses a centralised research database for international collaborative *in silico* clinical trials (**Chapter 5**). The system offers a secure data-exchange platform with high-quality datasets and trial protocols for multiple *in silico* planning comparisons of different treatment modalities. Quality assurance measures and automated data extraction procedures are applied for uniform analysis of the results.

Chapter 6 presents a ROCOCO (mistir.info/rococo) lung cancer trial that used the platform to investigate whether proton therapy can reduce dose to normal tissue, with equal or higher tumour dose. The trial participants downloaded de-identified clinical datasets to plan according to a strict protocol. The treatment planning results were returned to the server after which central analysis was performed. It showed that dose reduction is indeed possible, even when escalating the dose to the tumour.

Chapter 7 evaluates an online proton therapy decision support system (PRODECIS: prodecis.nl), reusing head and neck cancer data from another ROCOCO trial. The system assesses healthcare reimbursement eligibility for model-based tumour groups by comparison of photon and proton treatment plans, using publicly available toxicity and cost-prediction models from PredictCancer (predictcancer.org). The system proved successful in the assessment of 92% of the test cases and showed that, with proton therapy, 91% of the patients clinically benefit from reduced complications after one year, while 35% would be considered cost-effective.

Samenvatting

Radiotherapie-oncologie, sterker nog, de gezondheidszorg in z'n algemeen zit in de problemen. De uitgaven aan de gezondheidszorg stijgen snel en leiden tot een onhoudbare situatie [1]. Een vergrijzende bevolking verklaart de kosten ten dele, maar ook overbehandeling speelt een grote rol [2]. Dit kan enerzijds verklaard worden door defensieve geneeskunde [3], maar aan de andere kant ook door inefficiëntie ten gevolge van de langlevende klinische protocollen en de snelle *technology-push* vanuit de gezondheidsindustrie.

Een van de opties om kosten te reduceren, is door kosteneffectiviteitsanalyses of *Healthcare Technology Assessments* toe te passen om de politiek en besluitvormers te informeren over het juiste gebruik van (nieuwe) technologie.

Hiertoe is onderzocht hoe gepersonaliseerde beslishulpssystemen kunnen worden ontwikkeld, inclusief kosteneffectiviteit op het gebied van proton therapie. De volgende hypothesen zijn behandeld en bevestigd:

- ◆ *Rapid Learning Healthcare* (RLHC) kan klinische onderzoek en klinische besluitvorming bevorderen.
- ◆ De vooruitgang in *Healthcare Information Technology* (HIT) maakt het mogelijk om data internationaal te delen, terwijl de privacy van de patiënt gewaarborgd blijft.
- ◆ Het combineren van *in silico* planningsvergelijkingsstudies met multifactoriële predictiemodellen maakt het mogelijk om beperkte, dure middelen in te zetten, zoals geïntroduceerd bij protontherapie.

Het generieke concept van deze thesis is het gebruik van RLHC om *evidence-based medicine* aan te vullen met predictiemodellen om zo te voorzien in klinische beslishulpmiddelen voor een geïndividualiseerde behandeling van kanker (**Hoofdstuk 2**). RLHC past geavanceerde HIT toe om gescheiden klinische informatiebronnen samen te brengen in een datawarehouse dat een compleet, geïntegreerd, ziekte-georiënteerd overzicht geeft van medische gegevens voor onderzoeksdoeleinden. Dit hoofdstuk geeft een algemeen overzicht van enkele technieken die werden gebruikt in de daaropvolgende werken.

Hoofdstuk 3 demonstreert hoe instituten voordeel kunnen hebben van deze verbeterde presentatie van patiëntinformatie om de efficiency en kwaliteit van de dataverzameling voor klinische trials te verbeteren. Verder gebruikt RLHC *machine learning* principes om data te doorzoeken naar relevante relaties om zo, op data gebaseerde, complicatie voorspellingen te modelleren.

Voor het modelleren en valideren van deze modellen zijn grote hoeveelheden data nodig welke niet beschikbaar zijn binnen één enkel instituut. Verder is het noodzakelijk dat externe data beschikbaar zijn om de heterogeniteit van de data te vergroten wat de voorspellende waarde van de modellen vergroot en om externe validatie mogelijk te maken.

Hoofdstuk 4 presenteert de benodigde technieken om data te delen die de medische data silo's ontsluiten op een manier dat de patiënt privacy gewaarborgd wordt. Hiertoe worden 'connectoren' geïnstalleerd binnen de beveiligde netwerkomgevingen van het ziekenhuis. Door de data te transformeren in internationaal gestandaardiseerde, *semantisch interoperabele* datamodellen zijn computers in staat om data om te zetten in kennis. Verder worden medische gegevens vanuit de lokale taal vertaald in unieke, internationale codes wat betekent dat de data kunnen worden hergebruikt in het buitenland met gelijke betekenis en waarde. Dit maakt het mogelijk om op gefedereerde of gedistribueerde wijze complicatie predictiemodellen te leren [4].

Vooruitlopend op een gefedereerd netwerk, hetgeen erg geavanceerd en nog niet gemeengoed is, gebruikt het MISTIR (mistir.info) platform een gecentraliseerde onderzoek database om internationaal samen te werken aan *in silico* klinische trials (**Hoofdstuk 5**). Het systeem biedt een beveiligd data-uitwisselingsplatform met hoogwaardige datasets en studieprotocollen voor meerdere *in silico* planningsvergelijkingsstudies van verschillende behandelmodaliteiten. Voor een uniforme analyse van de resultaten worden kwaliteitsgarantiemaatregelen en automatische dataextracties toegepast.

Hoofdstuk 6 beschrijft een ROCOCO (mistir.info/rococo) longkanker studie die het platform gebruikt om te onderzoeken of protontherapie de dosis op het normale weefsel kan verminderen bij gelijke of hogere dosis op de tumor. De deelnemende centra downloaden geïdentificeerde klinische datasets om volgens een strikt protocol de behandelplannen te maken en vervolgens terug te sturen naar de centrale server waarna een centrale analyse werd gedaan. Dit onderzoek toonde aan dat stralingsreductie inderdaad mogelijk is, zelfs bij dosisescalatie.

In **Hoofdstuk 7** wordt een online protontherapie beslishulpsysteem geëvalueerd (PRODECIS, prodecis.nl), waarbij gegevens van een hoofd-hals ROCOCO studie werden hergebruikt. Het systeem beoordeelt of modelgebaseerde tumorgroepen in aanmerking komen voor een vergoeding van de ziektekostenverzekeraar door fotonen- en protonenplannen te vergelijken en gebruik te maken van openbaar aangeboden voorspellingsmodellen voor toxiciteit en kosten via PredictCancer (predictcancer.org). Het systeem bleek 92% van de testcasus-

sen goed te kunnen beoordelen en toonde aan dat 91% van de patiënten klinisch baat hebben bij protontherapie door verminderde complicaties na een jaar, terwijl 35% als kosteneffectief zou worden beschouwd.

References

1. CBS. Zorguitgaven stijgen met 1,8 procent in 2014. www.cbs.nl/nl-NL/menu/themas/gezondheid-welzijn/publicaties/artikelen/archief/2015/zorguitgaven-stijgen-met-1-8-procent-in-2014.htm [accessed November 8, 2015]
2. Patel TY. It Is Overtreatment, Not Overdiagnosis. *Acad Radiol* 2015;22(8):1044–1045. doi:10.1016/j.acra.2015.06.006
3. Sekhar Ms, Vyas N. Defensive medicine: A bane to healthcare. *Ann Med Health Sci Res* 2013;3(2):295. doi:10.4103/2141-9248.113688
4. Damiani A, Vallati M, Gatta R, Dinapoli N, Jochems A, Deist T, *et al.* Distributed Learning to Protect Privacy in Multi-centric Clinical Studies. In: Holmes JH, Bellazzi R, Sacchi L, Peek N, editors. *Artificial Intelligence in Medicine*, Springer International Publishing; 2015