From Data to Decision

A knowledge engineering approach
to individualise cancer therapy

Erik Roelofs

Cover
The Effect of a Drop: the foreground shows a combination of the W3C logo with the three sides of the tri-colour cube evoking the triplet of the RDF model. The peeled back lid invites you to *Open Your Data* to the Semantic Web. Projected on top are the recycling arrows that urge us to *Reuse* our data, to *Reduce* the occurrence of cancer and the healthcare costs involved by *Recycling* data into knowledge and apply that to our processes in a continuous, reviving loop depicted by the encompassing Ouroboros snake. The background shows the *ripples* extending on still water after a drop falls down and bounces back to do it again. Now imagine the *shockwaves* you get when all our data continuously and consistently drops in all those data-lakes out there!

Are you ready for a ride ?

From Data to Decision

A knowledge engineering approach to individualise cancer therapy

Hendrik Ambrosius Roelofs
iD 0000- 0003-2172-8669

Promotores
Prof. dr. P. Lambin
Prof. dr. ir. A.L.A.J. Dekker

Beoordelingscommissie
Prof. dr. G.G. (Frits) van Merode (voorzitter)
Prof. dr. P. Coucke
Prof. dr. W.J. Niessen
Prof. dr. J.E. Wildberger

– Scientia potentia est –

Voor Daniëlle,
Karsten en Imke

Contents

Chapter 1

Introduction

Treatment of choice

According to the World Health Organisation's estimations, 8.2 million people died from cancer in 2012 [1]. For more than 50% of cancer patients a form of radiation therapy is indicated [2]. With the increasing arsenal of possible treatment options (e.g. beam shapes, energy ranges, radiation modalities) and large patient heterogeneity it is getting ever harder to decide on the best possible treatment for an individual patient. Providing evidence of superiority of one treatment option over another is generally done by assessment in *clinical trials*. Under the same conditions, two treatment options are evaluated for differences in effectiveness and outcome. Most healthcare systems require the highest –level I– type of evidence via a randomised controlled trial (RCT) before reimbursing a treatment.

However, for particle therapy, there is still a lack of such evidence as not many RCTs have been performed or are not finished yet. Many have debated whether performing RCTs is a proper approach for particle therapy using the argument of a lack for clinical equipoise as there is a clear benefit regarding the physical dose distribution when compared to photon therapy. However, because the costs of particle therapy are notably higher, some argue that RCTs should be performed first to answer this question [3–13].

In this case, it is not clear whether one should adopt the new technology, risking investment in a suboptimal therapy, or should wait for more evidence, risking that patients do not receive on optimal treatment. This choice is particularly difficult with particle therapy, because the choice for adoption comes with high costs that cannot be reverted. It was shown for inoperable stage I non-small cell lung cancer that the preferred option for the Netherlands is to adopt the proton therapy and perform trials with that new technology instead of waiting for evidence first [14].

As will be shown in the next section, the use of models to predict the outcome of a treatment is well-developed and common practice in radiotherapy. By feeding simulated treatment planning results into these models, it is possible to predict the treatment outcome of both photon and particle therapy, without the patients undergoing the real treatment. This concept is used to simulate clinical trials and therefore referred to as *in silico*[1] *trials*.

---

[1] Per computer simulation, analogous to *in vitro* or *in vivo*

Part of the work presented here has contributed to a unique situation where the Dutch government and health insurance companies have accepted to adopt this *in silico* planning comparison principle and not only approve the treatment of standard proton therapy indications, but also the treatment of patients with proton therapy for certain less common indications with the obligation to perform prospective cohort studies. For four *model-based indications* (breast, head and neck, lung and prostate cancer), it is needed to test upfront whether there is an expected dosimetric benefit and whether this leads to a significant reduction of complications. Only then the treatment will be reimbursed; this is called the *model-based* approach [15,16].

Prediction modelling

In the treatment of cancer, a balance must be made between controlling the tumour growth and possible side effects or complications due to the treatment. When performing surgery, this mainly plays a local role, whereas for chemotherapy this concerns the whole body. For radiotherapy, the irradiation not only kills tumorous cells but normal ones that are within the fields as well. This is, however, a typically regional problem that is addressed during careful treatment planning. Moreover, a certain amount of cell-kill is accounted for as the human body often can restore cells of which the DNA is not beyond repair. And there is a clear dose-effect relationship, for acute as well as long-term side effects.

The use of models to describe the relation between treatment and its outcome is longstanding practice in radiotherapy. Typically, these are depending on given (or estimated) doses from the treatment planning system (TPS), although more recently models that use different metrics, such as patient characteristics, genomics and imaging features, as input are getting more common.

Dose-driven models

The research on the biologic effects of cell irradiation, known as *radiobiology*, has a long history. Soon after Röntgen's discovery of X-rays (1895) and radium by the Curies (1898) a flourishing era of radiation research and therapy began. Research in the first decades of the 20th century on animals and humans formed the basis of radiobiological theories on fractionation and dose-response relations that still hold in current times [17,18]. Considering the fact that cell-kill is a stochastic (probabilistic) effect – i.e. an increase in dose will increase the likelihood
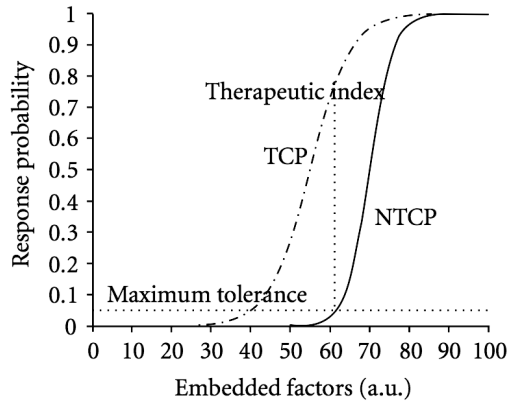
of kill – models describing the curative effect on the tumour and the damaging on the normal tissue are called Tumour Control Probability (TCP) and Normal Tissues Complication Probability (NTCP) models (Figure 1).

Most known TCP and NTCP models are dose-driven, which makes them easy to apply in clinical practice with the dose coming from the TPS simulations [20]. Although some vendors have included these models into their TPSs or released them as standalone products, their use is

Figure 1 Sigmoidal curves represent tumour control probability (TCP) and normal tissue complication probability (NTCP) as a function of treatment factors. The probabilities could be constructed as a function of heterogeneous variables (dose-volume metrics, biomarkers, and clinical factors). The radiotherapy treatment objective is to maximize the therapeutic index for each patient case [19].

not widespread. One of the reasons that prevented the uptake is that there is considerable doubt about the robustness of the models due to the uncertainty of the parameter values driving them. When used, it is mostly for the comparison of different treatment plans – of the same patient – relatively against each other or within the plan optimisation algorithms.

Another reason is that there are many other known and unknown confounding factors that influence treatment outcome. For instance, for lung cancer it is known that the mean lung dose (MLD) or the lung volume that receives 20 Gy are strong predictors for radiation-induced side effects such as pneumonitis [21–23]. However, it is more important to take patient-specific characteristics into account such as lung function, smoking status, tumour stage or volume [24]. When considering the dose-response curves schematically depicted in Figure 1, these factors shift or alter the relation such that with the same dose a different outcome is given. It is the inclusion of these patient-specific factors that is needed for personalised or *individualised healthcare*.

Bear in mind: physicians have always applied personalised medicine by taking into account personal factors such as age, family conditions or clinical history [25]. However, most of the current guidelines and tools to choose the best treatment for a patient are based on clinical trials. Considering the fact that only 3-5% of cancer patients are enrolled in clinical trials [26–29] and knowing that trial inclusion is typically under fairly optimal conditions (limited age, favoura-

ble health status, etc.) chances are high that the patient in front of the doctor is not within that trial group and thus will not receive an optimal individual treatment. Moving away from the "one model fits all" approach towards patient-specific treatment outcome prediction requires a different type of model generation than by means of regular clinical trials.
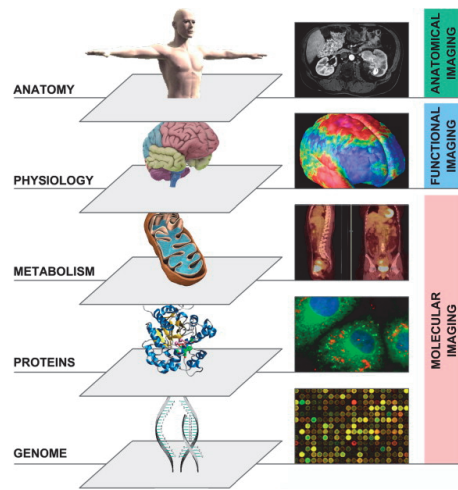
Data-driven models

With the advances in the discovery and application of innovative techniques such as genome sequencing, systems biology and diagnostic imaging, patients are more "transparent" than ever. Genome sequencing, for instance, discovers the exact sequence of nucleotide bases in the human DNA. New mathematical, holistic systems biology techniques offer ways to discover the underlying cell-signalling networks, understand how these contribute to diseases and indicate ways to cure these. Furthermore, continuous improvements in diagnostic imaging devices allow us to

Figure 2 Multilevel imaging: anatomical, functional and molecular imaging [48].

visualise a patient from a macroscopic level of body size, structure and organ functionality towards microscopic levels of cell type and activity through CT, MRI, ultrasound (US) and PET with ever increasing resolution [30].

The sheer amount and diversity of this multi-modal information as shown in Figure 2 illustrates the complexity of the problem to comprehend these mutual influences and relations to radiotherapy response. Discovery of these interactions through regular investigation with clinical trials will be too slow, if not impossible at all. While they provide data of high quality, trials are targeted to prove theories with typically singular endpoints amongst otherwise similar conditions. Investigating a plurality of unknown relations requires other means.

Knowledge engineering

With increasing digitisation and availability of medical data through recent developments and vast investments in healthcare information technology such

as electronic medical records (EMR), an effective method of modelling and relatively new in healthcare is possible by means of data mining and *machine learning* techniques. Using historical and daily clinical data, new statistical models can be trained to predict treatment outcome such as survival, toxicity and quality of life [31–35].

The field of research, which encompasses data mining, machine learning and "Big Data" storage techniques, is also called *Knowledge Engineering* [36,37]. It indicates that the key effort is to craft knowledge from large amounts of data, which is very difficult if not impossible to do by human interpretation. By finding patterns in the data through sorting and clustering and subsequently fitting of algorithms, it is possible to form models that, for instance, predict the chance that a patient will suffer from severe toxicity by a suggested treatment. Examples of such models are available online at eortc.be/ tools/gbmcalculator/, research.nki.nl/ibr and predictcancer.org/.

The creation of models by the extraction, fitting and modelling thus requires large amounts of data. Furthermore, because such models need proper validation, more datasets are needed from independent, preferably external data sources [38,39]. Therefore, adequate management of research or trial data as standard component in daily clinical routine with integration in the EMR is essential [40,41].

Data management

For the prediction of treatment outcome, the *power* and usability of a model depend on the quality of the data that was used during its generation. With respect to modelling through machine learning techniques, it is said that a lack of data quality can be compensated by increased data quantity. While often rebutted by the saying "garbage in, garbage out", there is truth in both statements. One false data entry will be statistically insignificant among thousands of others, unless it is a scaling factor, for instance. Or when a dataset is incorrectly labelled or linked to some score, it may be many but wrong data.

Data quality

Typically, historical clinical data is not well defined, although the increased use of Electronic Medical Records (EMR) in radiotherapy has structured the way

patient information is stored. Because data are often entered in free-text, they are rarely of the same quality as that from prospective randomised clinical trials.

Improving quality of data capture during treatment to complement trial data should be possible by increasing user-friendliness of EMRs [42] and by better compliance when entering information. Furthermore, it calls for good *data governance*, which emphasizes the *responsibility* of an organisation to adopt a thorough Quality Assure (QA) programme that encompasses Quality Control (QC) procedures, definitions of ownership and policies for (re-)use in both care as well as research.

When considering the quality of images, routinely performed diagnostic CT imaging is well established, especially in radiotherapy, for dose calculations depend on it. Due to international standards and calibration procedures, these data can be used consistently for data mining of images, for instance. However, although PET and MRI imaging is available for decades, its use in radiotherapy is not on par with CT imaging. One effect hereof is that quantitative standardisation is not as well accomplished. Recently, efforts have been made to harmonise calibration and QA procedures of PET and MRI as well [43–46].

Data quantity

Patients characteristics are very heterogeneous in a genetically, physically or even environmentally sense. This means that prediction models should be able to predict outcome for a large range of variables. In other words, there is a large multidimensional search or solution space for which the model should be valid. Typically, prediction models are optimised for interpolation within this space, not extrapolation. Furthermore, the use of predictions or related treatment choices that lie outside the known boundaries is usually not accepted in evidence-based medicine. To fit a wide range of patients, it is therefore essential to have access to plentiful, rich datasets. Conversely, with increasing individualisation of healthcare and increased information resolution, there is the desire to search for relevance among smaller differences in treatment response. An increased amount of data is needed to detect and predict these small differences.

One specific development that contributes to a wealth of patient data suitable for data mining and machine learning is the field known as *radiomics*, referring to the application of *genomics* and other 'omics' principles in radiotherapy and radiology. Using the anatomical and functional images as from the top layer in Figure 2, a diversity of shape and texture features is extracted and classified. Next, by applying techniques known from high-volume, high-throughput microarray filtering and clustering (Figure 3), the features can be analysed and used for outcome prediction modelling [47–49].

In contrast to biopsy-based molecular essays, this non-invasive "fingerprinting" technique has the benefit of taking spatial and temporal tumour heterogeneity into account for modelling treatment outcome.

Figure 3 Unsupervised hierarchical clustering of lung tumor image features extracted from CT images from 276 non-small cell lung cancer patients [50].

Data warehousing

Whether collected from clinical routine or for clinical trials, it is often difficult to get hold of relevant and correct data for machine learning. Trial data are frequently gathered by hand from scattered sources, which is prone to errors and time-consuming (**Chapter 3**). Any paper or film material needs to be digitised or is considered too old, although it could still be relevant as radiation effects may present themselves late and it would add to a more heterogeneous data pool. Research data is typically stored in disparate spreadsheets or databases throughout the institution or even stored at some external organisation that is leading the trial.

Patient data in radiotherapy institutions consists of several components and is stored in disparate information systems:

- the EMR for demographics and medical data,
- the TPS with CTs, contours, plans and dose calculations,

- the record and verify (R&V) system, which verifies planned and given doses and

- the picture archiving and communication system (PACS), which initially stored medical images only, but nowadays often stores exported data from TPSs as well.

By using data warehouse (DWH) techniques that centrally collect information from these different databases, using extraction, transformation and loading (ETL) pipelines, pruned data can be offered in one single environment for further research (**Chapter 3**, [50,51]). A core component is the transformation where a "view" is created of the data that is appropriate for further use by researchers. For instance, demographics data are filtered and patients that object the use of their data for research are excluded. Different views can be created as query results or via a user interface to facilitate multiple trials or investigations.

Another common practice is to transform the patient-centric data into a disease-centric presentation, as this is more in line with the way research is performed. Care must be taken that when one of the sources changes its data definition or structure, the DWH is updated through versioning to correctly capture the new information. One example is when international standards are updated, such as Common Toxicity Criteria (CTC). Instead of changing CTC scores at the original locations, it is best to change ETL queries to include transformation from one scoring version to the other and offer uniform data to the end-users, while ensuring backwards compatibility.

With the data being readily available in databases, it is now possible to add data extraction services that generate new information from existing data. In **Chapter 3**, it is shown that previously uncalculated DVH metrics can be generated automatically and improve data quality and collection time for trials. Furthermore, it facilitates the initiation of new trials as patients can easily be selected for inclusion.

Data sharing

In **Chapter 2**, the application of *Rapid Learning* in the realm of radiotherapy is described. While originally referring to fast learning methods in the sense of teaching by electronic means (e-learning), Rapid Learning here refers to frameworks of Healthcare Information Technology (IT) systems (such as EMR, DWH, etc.) and shared procedures to exchange information among the healthcare community in general and cancer care in particular [52–54]. Given

the fact that data from institutions are insufficient to feed learning and prediction systems locally, a large effort is needed in crossing institutional and national borders to share that information among fellow research centres.

Issues that arise when sharing information among healthcare centres are numerous [55,56]. First of all, there is a lack of willingness to share as research data have long been seen as precious and essential for an institution's status. Indeed, the data collection is typically time-consuming and offers a basis for publications and therefore has its (political) value. However, the value of such individual collections is diminishing when international data pooling is considered for prediction modelling, for example [57]. It is increasingly accepted that not the data itself but the extracted knowledge is of future interest and importance.

Second, with the frequent reporting of large-scale unauthorised monitoring of private information, there are sincere concerns of the public about sharing their medical data, which can cause people to ask for exclusion of trials and data-sharing initiatives. National privacy laws are not always easy to understand or implement, with a huge increase in difficulty and needed effort when considering international jurisdiction [55]. Whether justified or not, the scepticism for inadequate capacity of the governmental coordination of an infrastructure to link medical records countrywide, has caused the Dutch senate to cancel the initial efforts for a national virtual EMR after 15 years of work. The main concerns were related to the (coordination of) ownership of the data and privacy [58].

Thirdly, from a technical point of view, it is not common practice yet for healthcare institutions to have full DWH solutions in place. Daily healthcare data is primarily stored due to the legal obligation of keeping a track record for 15 years. Reuse of that data is typically only after retrieving from archives for clinical trials. While using historical data for optimising the organisations processes and workflow is common in industry, this is not yet widely incorporated into the field of healthcare. This is changing, however. For instance, while previously PACS vendors where targeting business solutions for storing images alone, there is now a growing industrial effort to expand their reach towards any available data with infrastructures such as Enterprise Content Management (ECM) and Vendor-Neutral Archive (VNA) systems. To properly organise this, the Integrating the Healthcare Enterprise (IHE) organisation, whose goal is to improve interoperability in healthcare, has introduced a new profile for Cross-Enterprise Document Workflow (XDW) [59,60].

Common data language

This leads to an extremely challenging issue concluded with here: a common language. Literally, EMR data is most often entered in the national language. To exchange such information internationally, the data has to be translated. A typical situation is that a data manager is translating the information and manually entering the data into a central trial database. Large-scale data capture and exchange is not an option that way.

However, it is even more difficult that not always the same meaning is given to words or combinations thereof. A combination of computer systems that is able to communicate and exchange information for the purpose of machine learning is called *semantically interoperable* if no human intervention is needed to translate or give meaning to the data [61–64]. Based on a *syntactic* layer of standards for data format and transfer syntax, such as the Health Level Seven (HL7) standard for healthcare, the data's interpretation is given via metadata (data about the data) that links data elements to concepts such as defined in SNOMED Clinical Terminology (SNOMED-CT) in a chosen shared *ontology* [65,66]. With such an internationally defined standard, the relation and reasoning between concepts is determined, which makes it possible to model knowledge in a globally unique way.

Knowledge-driven healthcare

Aging and improved diagnostics has increased healthcare costs globally. One of the options to gain efficacy and efficiency is by incorporating advanced information technology solutions to match the right treatment to the right patient. Data warehousing and mining to provide data-driven prediction models and semantic interoperability to improve data sharing are some of the solutions as discussed above.

Extracting information and gaining knowledge that can be applied for decision-making in daily practice is what drives current innovations in healthcare. However, it is important to constantly update the information pool and revise findings. Although relying on data sharing as a means, it is essential to share the gained knowledge itself as well. Figure 4 shows how the concept of *knowledge-driven healthcare* can help improve radiotherapy outcome by feeding knowledge from previous follow-up data and regular clinical evidence back in a cascade setup to support decision-making and new evidence finding.

Knowledge-driven Healthcare
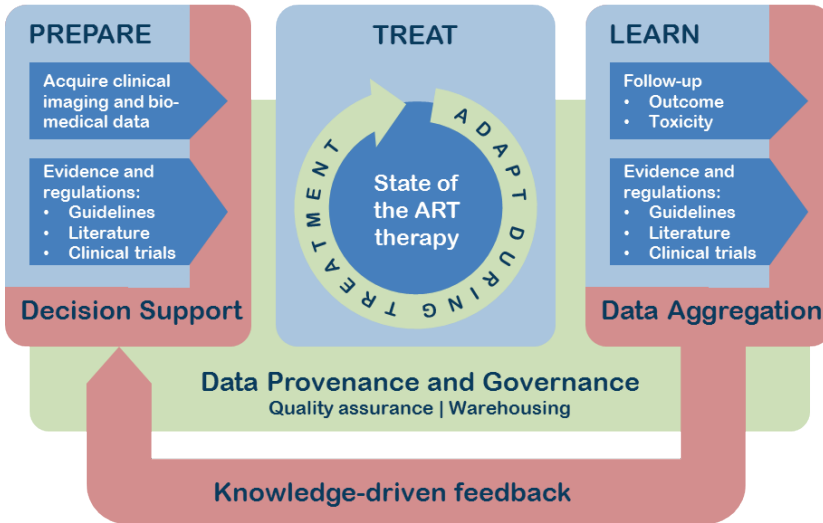
Figure 4 Knowledge-driven healthcare principles using a clinical decision support system in conjunction with standard evidence and regulations to choose the optimal treatment. Afterwards, by learning from follow-up data, that knowledge is fed back to improve the clinical decision support system and adapt regulations [70].

Multicentric in silico trials

An example of a centralised collaboration network is the Multicentric *In Silico* Trials In Radiotherapy (MISTIR: mistir.info) framework. In **Chapter 5**, it is shown how this facility was set up for the Radiation Oncology COllaborative COmparison (ROCOCO) trial, which simulates clinical trials that compare photon, proton and carbon-ion radiotherapy for lung, head and neck and prostate cancer. MISTIR offers the central database with clinical datasets and treatment planning protocols that are securely shared among the ROCOCO partners. Various QA methods are provided to assure proper data quality after transfer and accurate analyses. Currently, multiple ROCOCO trials are undertaken, including studies in lung, head and neck, prostate and recently neuro-oncology as well (mistir.info/activities).

Chapter 6 describes how the platform was used for the multicentric *in silico* comparison of the treatment of lung cancer patients with photon and proton therapy. The participants downloaded the datasets and performed treatment planning according to the provided protocol. The results were uploaded to the data centre, where dose-volume metrics from the different plan modalities were

extracted and analysed. The results were collaboratively discussed and finally published.

The advantage of the MISTIR setup is that the high-quality datasets and results are centrally stored and easily available for reuse and analysis in subsequent projects. One such a project was ENVISION, which was a collaboration of sixteen leading European research centres and industrial partners, coordinated by CERN [67]. Under the umbrella of ENLIGHT, the FP7[2] supported European Network for Light Ion Hadron Therapy, the project aimed at improving the quality assurance tools for hadron therapy by developing solutions for quantitative imaging, monitoring of delivered dose, adaptive treatment planning through measurements and simulation studies, among others. Part of the RO-COCO data was used for the investigation of photon, proton and carbon-ion plan quality due to organ motion [68,69].

Decision support systems

With the increase of patient-related information and treatment options, it is increasingly difficult for healthcare professionals to choose the best available treatment for patients. Knowledge-driven Healthcare (Figure 4) helps to achieve this by constantly updating decision-making guidelines based on (*in silico*) trials as well as daily clinical data (**Chapter 6**, [70]). Furthermore, by using data-driven models derived through collaborative efforts of shared data and knowledge networks, decision support tools can be generated to assist in choosing the right treatment and make the transition of population-based towards individualised medicine (**Chapter 2,** [71]).

The *model-based* approach that was introduced previously uses this combination of *in silico* planning comparisons and complication prediction models to assess the eligibility for proton therapy. In **Chapter 7,** a proton therapy decision support system is presented (PRODECIS: prodecis.nl). This system offers an online service that compares extracted dose-volume metrics from photon and proton treatment plans. After feeding these metrics into complication prediction algorithms, a distinction is made whether a patient is expected to benefit from proton therapy or not. The toxicity and cost-prediction models that were used are publicly available from PredictCancer (predictcancer.org).

---

[2] Seventh Framework Programmes for Research and Technological Development, European Union

References

1. Stewart BW, Wild C, International Agency for Research on Cancer, World Health Organization, editors. *World cancer report 2014*. Lyon, France: International Agency for Research on Cancer; 2014

2. Delaney G, Jacob S, Featherstone C, Barton M. The role of radiotherapy in cancer treatment. *Cancer* 2005;104(6):1129–1137. doi:10.1002/cncr.21324

3. Van Loon J, Grutters J, Macbeth F. Evaluation of novel radiotherapy technologies: what evidence is needed to assess their clinical and cost effectiveness, and how should we get it? *Lancet Oncol* 2012;13(4):e169–e177. doi:10.1016/S1470-2045(11)70379-5

4. Brada M, Pijls-Johannesma M, Ruysscher DD. Proton Therapy in Clinical Practice: Current Clinical Evidence. *J Clin Oncol* 2007;25(8):965–970. doi:10.1200/JCO.2006.10.0131

5. Glimelius B, Montelius A. Proton beam therapy - do we need the randomised trials and can we do them? *Radiother Oncol* 2007;83(2):105–109. doi:10.1016/j.radonc.2007.04.009

6. Lodge M, Pijls-Johannesma M, Stirk L, Munro AJ, De Ruysscher D, Jefferson T. A systematic literature review of the clinical and cost-effectiveness of hadron therapy in cancer. *Radiother Oncol* 2007;83(2):110–122. doi:10.1016/j.radonc.2007.04.007

7. Olsen DR, Bruland OS, Frykholm G, Norderhaug IN. Proton therapy - a systematic review of clinical effectiveness. *Radiother Oncol* 2007;83(2):123–132. doi:10.1016/j.radonc.2007.03.001

8. Goitein M, Cox JD. Should Randomized Clinical Trials Be Required for Proton Radiotherapy? *J Clin Oncol* 2008;26(2):175–176. doi:10.1200/JCO.2007.14.4329

9. Bentzen SM. Randomized Controlled Trials in Health Technology Assessment: Overkill or Overdue? *Radiother Oncol* 2008;86(2):142–147. doi:10.1016/j.radonc.2008.01.012

10. Suit H, Kooy H, Trofimov A, Farr J, Munzenrider J, DeLaney T, *et al.* Should positive phase III clinical trial data be required before proton beam therapy is more widely adopted? No. *Radiother Oncol* 2008;86(2):148–153. doi:10.1016/j.radonc.2007.12.024

11. Tepper JE. Protons and Parachutes. *J Clin Oncol* 2008;26(15):2436–2437. doi:10.1200/JCO.2008.17.1165

12. Zietman A, Goitein M, Tepper JE. Technology Evolution: Is It Survival of the Fittest? *J Clin Oncol* 2010;28(27):4275–4279. doi:10.1200/JCO.2010.29.4645

13. Sheehan M, Timlin C, Peach K, Binik A, Puthenparampil W, Lodge M, *et al.* Position statement on ethics, equipoise and research on charged particle radiation therapy. *J Med Ethics* 2014;40(8):572–575. doi:10.1136/medethics-2012-101290

14. Grutters JPC, Abrams KR, de Ruysscher D, Pijls-Johannesma M, Peters HJM, Beutner E, *et al.* When to wait for more evidence? Real options analysis in proton therapy. *The Oncologist* 2011;16(12):1752–1761. doi:10.1634/theoncologist.2011-0029

15. Langendijk JA, Lambin P, De Ruysscher D, Widder J, Bos M, Verheij M. Selection of patients for radio-therapy with protons aiming at reduction of side effects: the model-based approach. *Radiother Oncol* 2013;107(3):267–273. doi:10.1016/j.radonc.2013.05.007

16. Widder J, van der Schaaf A, Lambin P, Marijnen CAM, Pignol JP, Rasch CR, *et al.* The quest for evidence for proton therapy: the model-based approach and precision medicine. *Int J Radiat Oncol* 2015. doi:10.1016/j.ijrobp.2015.10.004

17. Joiner M, Kogel A van der, editors. *Basic clinical radiobiology*. 4th ed. London: Hodder Arnold; 2009

18. Willers H, Beck-Bornholdt HP. Origins of radiotherapy and radiobiology: separation of the influence of dose per fraction and overall treatment time on normal tissue damage by Reisner and Miescher in the 1930s. *Radiother Oncol* 1996;38(2):171–173. doi:10.1016/0167-8140(95)01656-2

19. Spencer SJ, Almiron Bonnin D, Deasy JO, Bradley JD, El Naqa I. Bioinformatics Methods for Learning Radiation-Induced Lung Inflammation from Heterogeneous Retrospective and Prospective Data. *J Biomed Biotechnol* vol. 2009, Article ID 892863, 14 pages, 2009. doi:10.1155/2009/892863

20. Li XA, Alber M, Deasy JO, Jackson A, Jee KWK, Marks LB, *et al.* The use and QA of biologically related models for treatment planning: Short report of the TG-166 of the therapy physics committee of the AAPM a). *Med Phys* 2012;39(3):1386–1409. doi:10.1118/1.3685447

21. Rodrigues G, Lock M, D'Souza D, Yu E, Van Dyk J. Prediction of radiation pneumonitis by dose - volume histogram parameters in lung cancer—a systematic review. *Radiother Oncol* 2004;71(2):127–138. doi:10.1016/j.radonc.2004.02.015

22. Palma DA, Senan S, Tsujino K, Barriger RB, Rengan R, Moreno M, *et al.* Predicting radiation pneumonitis after chemoradiation therapy for lung cancer: an international individual patient data meta-analysis. *Int J Radiat Oncol Biol Phys* 2013;85(2):444–450. doi:10.1016/j.ijrobp.2012.04.043

23. Herrmann PDT, Geyer P, Appold DS. The mean lung dose (MLD). *Strahlenther Onkol* 2015;191(7):557–565. doi:10.1007/s00066-015-0833-3

24. Colinet B, Jacot W, Bertrand D, Lacombe S, Bozonnat MC, Daurès JP, *et al.* A new simplified comorbidity score as a prognostic factor in non-small-cell lung cancer patients: description and comparison with the Charlson's index. *Br J Cancer* 2005;93(10):1098–1105. doi:10.1038/sj.bjc.6602836

25. Schilsky RL. Personalizing Cancer Care: American Society of Clinical Oncology Presidential Address 2009. *J Clin Oncol* 2009;27(23):3725–3730. doi:10.1200/JCO.2009.24.6827

26. Jenkins V. Discussing randomised clinical trials of cancer therapy: evaluation of a Cancer Research UK training programme. *BMJ* 2005;330(7488):400–0. doi:10.1136/bmj.38366.562685.8F

27. Grand MM, O'Brien PC. Obstacles to participation in randomised cancer clinical trials: a systematic review of the literature. *J Med Imaging Radiat Oncol* 2012;56(1):31–39. doi:10.1111/j.1754-9485.2011.02337.x

28. Murthy VH, Krumholz HM, Gross CP. Participation in cancer clinical trials: race-, sex-, and age-based disparities. *J Am Med Assoc* 2004;291(22):2720–2726. doi:10.1001/jama.291.22.2720

29. Movsas B, Moughan J, Owen J, Coia LR, Zelefsky MJ, Hanks G, *et al.* Who enrolls onto clinical oncology trials? A radiation Patterns Of Care Study analysis. *Int J Radiat Oncol Biol Phys* 2007;68(4):1145–1150. doi:10.1016/j.ijrobp.2007.01.051

30. Townsend DW. Multimodality imaging of structure and function. *Phys Med Biol* 2008;53(4):R1. doi:10.1088/0031-9155/53/4/R01

31. El Naqa I, Bradley J, Blanco AI, Lindsay PE, Vicic M, Hope A, *et al.* Multivariable modeling of radiotherapy outcomes, including dose–volume and clinical factors. *Int J Radiat Oncol • Biol • Phys* 2006;64(4):1275–1286. doi:10.1016/j.ijrobp.2005.11.022

32. Carvalho S, Leijenaar RTH, Velazquez ER, Oberije C, Parmar C, van Elmpt W, *et al.* Prognostic value of metabolic metrics extracted from baseline positron emission tomography images in non-small cell lung cancer. *Acta Oncol* 2013;52(7):1398–1404. doi:10.3109/0284186X.2013.812795

33. El Naqa I. Biomedical informatics and panomics for evidence-based radiation therapy. *Wiley Interdiscip Rev Data Min Knowl Discov* 2014;4(4):327–340. doi:10.1002/widm.1131

34. Van Soest JPA, Dekker ALAJ, Roelofs E, Nalbantov G. Application of Machine Learning for Multicenter Learning. In: Naqa IE, Li R, Murphy MJ, editors. *Machine Learning in Radiation Oncology*, Springer International Publishing; 2015

35. Robertson SP, Quon H, Kiess AP, Moore JA, Yang W, Cheng Z, *et al.* A data-mining framework for large scale analysis of dose-outcome relationships in a database of irradiated head and neck cancer patients. *Med Phys* 2015;42(7):4329–4337. doi:10.1118/1.4922686

36. Studer R, Benjamins VR, Fensel D. Knowledge engineering: Principles and methods. *Data Knowl Eng* 1998;25(1–2):161–197. doi:10.1016/S0169-023X(97)00056-6

37. Meyer AM, Basch E. Big Data Infrastructure for Cancer Outcomes Research: Implications for the Practicing Oncologist. *J Oncol Pract* 2015;11(3):207–208. doi:10.1200/JOP.2015.004432

38. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, *et al.* Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and ElaborationThe TRIPOD Statement: Explanation and Elaboration. *Ann Intern Med* 2015;162(1):W1–W73. doi:10.7326/M14-0698

39. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *The BMJ* 2015;350:g7594. doi:10.1136/bmj.g7594

40. Schreiweis B, Trinczek B, Köpcke F, Leusch T, Majeed RW, Wenk J, *et al.* Comparison of electronic health record system functionalities to support the patient recruitment process in clinical trials. *Int J Med Inf* 2014;83(11):860–868. doi:10.1016/j.ijmedinf.2014.08.005

41. Gupta SK. Paperless clinical trials: Myth or reality? *Indian J Pharmacol* 2015;47(4):349–353. doi:10.4103/0253-7613.161247

42. Ratwani RM, Fairbanks RJ, Hettinger AZ, Benda NC. Electronic health record usability: analysis of the user-centered design processes of eleven electronic health record vendors. *J Am Med Inform Assoc JA-MIA* 2015. doi:10.1093/jamia/ocv050

43. Hatt M, Cheze le Rest C, Descourt P, Dekker A, De Ruysscher D, Oellers M, *et al.* Accurate Automatic Delineation of Heterogeneous Functional Volumes in Positron Emission Tomography for Oncology Applications. *Int J Radiat Oncol* 2010;77(1):301–308. doi:10.1016/j.ijrobp.2009.08.018

44. Levy MA, Freymann JB, Kirby JS, Fedorov A, Fennessy FM, Eschrich SA, *et al.* Informatics methods to enable sharing of quantitative imaging research data. *Magn Reson Imaging* 2012;30(9):1249–1256. doi:10.1016/j.mri.2012.04.007

45. Boellaard R. Need for standardization of 18F-FDG PET/CT for treatment response assessments. *J Nucl Med Off Publ Soc Nucl Med* 2011;52 Suppl 2:93S – 100S. doi:10.2967/jnumed.110.085662

46. Boellaard R, Delgado-Bolton R, Oyen WJG, Giammarile F, Tatsch K, Eschner W, *et al.* FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging* 2015;42(2):328–354. doi:10.1007/s00259-014-2961-x

47. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, *et al.* Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;48(4):441–446. doi:10.1016/j.ejca.2011.11.036

48. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Cavalho S, *et al.* Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014;5. doi:10.1038/ncomms5006

49. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, *et al.* Radiomics: the process and the challenges. *Magn Reson Imaging* 2012;30(9):1234–1248. doi:10.1016/j.mri.2012.06.010

50. Wisniewski MF, Kieszkowski P, Zagorski BM, Trick WE, Sommers M, Weinstein RA. Development of a Clinical Data Warehouse for Hospital Infection Control. *J Am Med Inform Assoc* 2003;10(5):454–462. doi:10.1197/jamia.M1299

51. El Fadly A, Rance B, Lucas N, Mead C, Chatellier G, Lastic PY, *et al.* Integrating clinical research with the Healthcare Enterprise: from the RE-USE project to the EHR4CR platform. *J Biomed Inform* 2011;44 Suppl 1:S94–S102. doi:10.1016/j.jbi.2011.07.007

52. Etheredge LM. A Rapid-Learning Health System. *Health Aff (Millwood)* 2007;26(2):w107–w118. doi:10.1377/hlthaff.26.2.w107

53. Friedman CP, Wong AK, Blumenthal D. Achieving a Nationwide Learning Health System. *Sci Transl Med* 2010;2(57):57cm29–cm57cm29. doi:10.1126/scitranslmed.3001456

54. Abernethy AP, Etheredge LM, Ganz PA, Wallace P, German RR, Neti C, *et al.* Rapid-Learning System for Cancer Care. *J Clin Oncol* 2010;28(27):4268–4274. doi:10.1200/JCO.2010.28.5478

55. EUROCRAN. Eurocan plus report: feasibility study for coordination of national cancer research activities. *Ecancermedicalscience* 2008;2:84. doi:10.3332/ecancer.2011.84

56. Sullivan R. Policy challenges for cancer research: a call to arms. *Ecancermedicalscience* 2007;1. doi:10.3332/ecancer.2008.53

57. Deasy JO, Bentzen SM, Jackson A, Ten Haken RK, Yorke ED, Constine LS, *et al.* Improving Normal Tissue Complication Probability Models: The Need to Adopt a "Data-Pooling" Culture. *Int J Radiat Oncol* 2010;76(3, Supplement):S151–S154. doi:10.1016/j.ijrobp.2009.06.094

58. Smits J. The future of the national virtual EPR in the Netherlands: Changes in approaching large-scale exchange of information in Dutch Healthcare. *J Chain-Comput* 2013;4. URN:NBN:NL:UI:10-1-114602

59. Noumeir R. Integrating the healthcare enterprise process. *Int J Healthc Technol Manag* 2008;9(2):167. doi:10.1504/IJHTM.2008.017371

60. IHE. *Cross-Enterprise Document Workflow (XDW) - Trial Implementation*. 2014

61. Garde S, Knaup P, Hovenga E, Heard S. Towards semantic interoperability for electronic health records. *Methods Inf Med* 2007;46(3):332–343. doi:10.1160/ME5001

62. Ibrahim A, Bucur A, Perez-Rey D, Alonso E, de Hoog M, Dekker A, *et al.* Case Study for Integration of an Oncology Clinical Site in a Semantic Interoperability Solution based on HL7 v3 and SNOMED-CT: Data Transformation Needs. *AMIA Jt Summits Transl Sci Proc AMIA Summit Transl Sci* 2015;2015:71

63. Rubin DL, Napel S. Imaging Informatics: Toward Capturing and Processing Semantic Information in Radiology Images. *IMIA Yearb 2010 Biomed Inform Build Capacity Worldw* 2010:34–42

64. Van Soest J, Lustberg T, Grittner D, Marshall MS, Persoon L, Nijsten B, *et al.* Towards a semantic PACS: Using Semantic Web technology to represent imaging data. *Stud Health Technol Inform* 2014:166–170. doi:10.3233/978-1-61499-432-9-166

65. SNOMED CT - The Global Language of Healthcare. www.ihtsdo.org/snomed-ct [accessed November 8, 2015]

66. Health Level Seven. www.hl7.org [accessed November 8, 2015]

67. European NoVel Imaging Systems for ION therapy. envision.web.cern.ch/ENVISION [accessed November 8, 2015]

68. Walsh S, Roelofs E, Verhaegen F. Comparing geometrical plan robustness and volatility of TCP for the ROCOCO photon prostate dataset. *Phys Med* 2013;29(5):571. doi:10.1016/j.ejmp.2013.05.021

69. Walsh S, Roelofs E, Kuess P, Lambin P, Jones B, Georg D, *et al.* A validated tumor control probability model based on a meta-analysis of low, intermediate, and high-risk prostate cancer patients treated by photon, proton, or carbon-ion radiotherapy. *Med Phys* 2016;43(2):734. doi:10.1118/1.4939260

70. Lambin P, van Stiphout RGPM, Starmans MHW, Rios-Velazquez E, Nalbantov G, Aerts HJWL, *et al.* Predicting outcomes in radiation oncology—multifactorial decision support systems. *Nat Rev Clin Oncol* 2013;10(1):27–40. doi:10.1038/nrclinonc.2012.196

71. Dekker A. Rapid Learning Approach for Decision Support Systems. *Phys Med* 2012;28(4):333. doi:10.1016/j.ejmp.2012.06.009

Chapter 2

INDIVIDUALISATION THROUGH *RLHC*

*Rapid Learning Health Care in Oncology –
an approach towards decision support
systems enabling customised radiotherapy*

Philippe Lambin, Erik Roelofs, Bart Reymen, Emmanuel Rios Velazquez,
Jeroen Buijsen, Catharina M.L. Zegers, Sara Carvalho, Ralph T.H. Leijenaar,
Georgi Nalbantov, Cary Oberije, M. Scott Marshall, Frank Hoebers,
Esther G.C. Troost, Ruud G.P.M. van Stiphout, Wouter van Elmpt,
Trudy van der Weijden, Liesbeth Boersma, Vincenzo Valentini, Andre Dekker

Abstract

Purpose: An overview of the Rapid Learning methodology, its results, and the potential impact on radiotherapy.

Material and results: Rapid Learning methodology is divided into four phases. In the data phase, diverse data are collected about past patients, treatments used, and outcomes. Innovative information technologies that support semantic interoperability enable distributed learning and data sharing without additional burden on health care professionals and without the need for data to leave the hospital. In the knowledge phase, prediction models are developed for new data and treatment outcomes by applying machine learning methods to data. In the application phase, this knowledge is applied in clinical practice via novel decision support systems or via extensions of existing models such as Tumour Control Probability models. In the evaluation phase, the predictability of treatment outcomes allows the new knowledge to be evaluated by comparing predicted and actual outcomes.

Conclusion: Personalised or tailored cancer therapy ensures not only that patients receive an optimal treatment, but also that the right resources are being used for the right patients. Rapid Learning approaches combined with evidence based medicine are expected to improve the predictability of outcome and radiotherapy is the ideal field to study the value of Rapid Learning. The next step will be to include patient preferences in the decision making.

Tailored cancer therapies, in which specific information about patients and tumours is taken into account during treatment decisions, are an important step forward from current population-based therapy [1]. However, given the developments outlined below, it is becoming increasingly difficult to identify the best treatment for an individual cancer patient:

- *Tumours and patients seem to be even less homogeneous than previously assumed,* meaning the same treatments can have different outcomes in patients who have the same type of tumour. For instance, there are at least four molecular subtypes of breast cancer, each with very different outcomes [2]. Based on gene signatures various subgroups of tumours can be identified [3–8].

- *The number of treatment options is increasing.* For example, early stage prostate cancer can now be treated with conservative treatment, prostatectomy, external radiotherapy, stereotactic radiotherapy, LDR or HDR brachytherapy, high-intensity focused ultrasound, hormone therapy, combination therapies and so on. A different example is the recent rise of targeted therapies that are rapidly growing in numbers. Performing classic randomised trials to compare all new treatment options with the "gold standard" is becoming impossible by the current speed of innovation.

- *The evidence for the right choice in an individual patient is inadequate.* First, 'evidence-based medicine' and the ensuing guidelines always lag somewhat behind practice, particularly in highly technological, innovative and rapidly evolving fields such as radiotherapy. In addition, translating the results of clinical trials to the general patient population and environment is not straightforward, given the higher quality of care in clinical trials and the known selection bias (trials reach no more than 3% of cancer patients, in radiotherapy this figure is even lower) [9–11]. Finally, given the developments mentioned above – more treatment options and less homogeneous patient groups – the urgency to scaffold our treatment decisions with robust knowledge and the demand for evidence-based medicine is larger than ever.

- *It is becoming more difficult to find the right evidence.* Despite – or perhaps due to – the fact that papers are being published in rapidly increasing numbers (e.g., as a radiation-oncologist specialising in lung cancer, has to read around eight articles per day to keep up with the literature

[12]), it is difficult to match the characteristics of the individual patient to evidence from the literature and to evaluate the quality of that evidence.
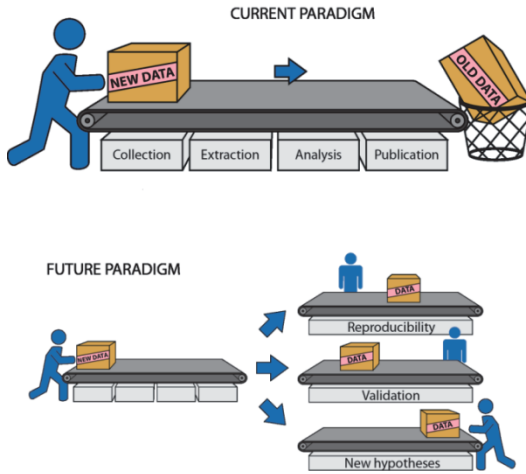

Figure 1 Current paradigm versus future paradigm

The developments illustrated above have given rise to a search for an alternative to the elaborate consensus and evidence-based guideline medicine format when it comes to making treatment decisions. The alternative discussed in this article is rapid learning [13]. Although it is known under various names, including Knowledge-driven Healthcare, Computer Assisted Theragnostics and Learning Intelligence Network, the basic idea in all cases is the (re)use of historical data from routine clinical practice for decisions concerning new patients or to test new hypothesis [14–19] (Fig. 1). This has a number of obvious advantages, such as the large number of readily available patients and less selection bias compared to clinical trials. However, it also has some important disadvantages; for example, the quality of the data in clinical practice is much lower than in clinical trials [20]. There is a long very successful history of putting genomic data public and reusing them [3–8].

This paper provides an overview of the methods used in Rapid Learning, the initial results, and an outlook as to how the techniques involved may influence clinical radiotherapy.

Methods and results

Rapid Learning involves four phases (Fig. 2) [13] which are continually iterated. In the data phase, data on past patients are collected, including
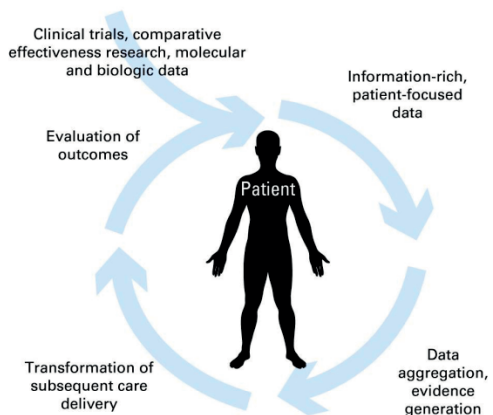

Figure 2 Four phases of rapid learning [13]

their delivered treatments and outcomes. In the knowledge phase, knowledge is generated from these data. In the application phase, this knowledge is applied to clinical practice. In the final evaluation phase, the outcomes are evaluated, after which the first phase starts again. In every phase, external knowledge (e.g., from clinical trials) is used to optimise the phase. The sections below describe the methods used and examples of typical results for every phase.

Data

Rapid Learning requires both a great deal of data and a large diversity of data. The amount of data is important (a) to obtain higher quality knowledge (the quality of the knowledge correlates with the number of patients on which that knowledge is based) and (b) to be able to generate knowledge concerning smaller, more homogeneous patient groups and/or use more variables in the knowledge phase. The diversity of the data (particularly with respect to the treatments used, but also in terms of patient characteristics) is important to ultimately decide which treatment is best for an individual patient.

Obtaining enough data of sufficient quality and diversity is the biggest challenge in Rapid Learning. This is only possible if data are shared across institutional and national borders, both academic and community health care systems. Such data sharing is hampered by a lack of time; differences in language and culture as well as data recording practices; the academic and political value of data; risks to reputation; privacy and legal aspects and so on. Nonetheless, one project that has made successful use of data sharing is euroCAT (www.eurocat.info), a collaborative project involving radiotherapy institutes in the Netherlands, Germany and Belgium. A crucial factor in the success of this project was the use of innovative information technologies, which made it possible to learn from each other's data without the data having to leave the institution (a concept known as distributed learning). Another important factor was the development of a dataset with semantic interoperability (also known as 'data with linguistic unity' or 'machine-readable data'), in which local terms are converted into concepts from a well-defined ontology (e.g., NCI Thesaurus, SNOMED). In such an approach, the ontology terms serve as a common interface to the data at each institutional site, enabling a common approach to information retrieval and reasoning facilitated through a semantic portal to the data. This semantic interoperability approach also allows one to add data from clinical trials to further strengthen the data available to Rapid Learning.

The data collected in routine clinical care are often of lower quality compared to data from clinical trials. Data captured in routine care are often incorrect, contradictory, missing and biased. Although many problems are mitigated by the sheer volume of data, it is important to include data quality improvement protocols varying from simple logic (e.g. it is impossible to be 60 kg and have a BMI of 32) to more probabilistic approaches (e.g. for a similar patient cohort the median value of the maximal standard uptake value from 18-FDG PET scans should be similar between two institutes). A positive effect of such initiatives is that they give rise to increasing coordination with respect to what data need to be collected and how (i.e., disease-specific 'umbrella' protocols). The end users of the knowledge, the provider and the patient, not only need to gain insight into effects of various treatment options, but also in uncertainties, conflicting data, and toxicities and other treatment burden.

It should be noted that getting data in the proposed manner does not mean that there is a need to capture more data, which would be an unacceptable additional burden to often overloaded professionals. Rather, the data that are already captured in routine care and in clinical trials are combined and re-used. There are various prototypes to do this such as in the euroCAT project where a fully automated, daily synchronisation of the clinical databases into a semantically interoperable dataset takes place.

Knowledge

Machine learning is used to extract knowledge from great amounts of data. In machine learning, models/algorithms are developed that best describe the data but that can also make predictions for new, unseen data. Models trained on retrospective data may be used to predict the outcomes (e.g., survival, quality of life, toxicity, etc.) of various treatments on the basis of data from a new patient. Obviously, it is crucial that such models are adequately validated [21]; an unvalidated model is of very limited value. To this end, a validation set should always be available, preferably from a different institute than that from which the data were used to create the model. Examples of radiotherapy models (on the basis of both clinical trials and Rapid Learning) are available for non-small cell lung, rectal and head-and-neck cancer on www.predictcancer.org, breast cancer on research.nki.nl/ibr and glioblastoma on www.eortc.be/tools/gbmcalculator.

Application

In this phase, the knowledge generated by Rapid Learning is applied with the help of decision support systems (DSS). Typically, these are tools and software applications that can be used to apply knowledge-driven healthcare in practice. Examples include nomograms (as in Fig. 3) [14,15,22–26] and websites such as those named above, for radiotherapy models, which help predict the expected treatment outcome of radiotherapy when they are supplied with the parameters specifically relevant to the clinical case.
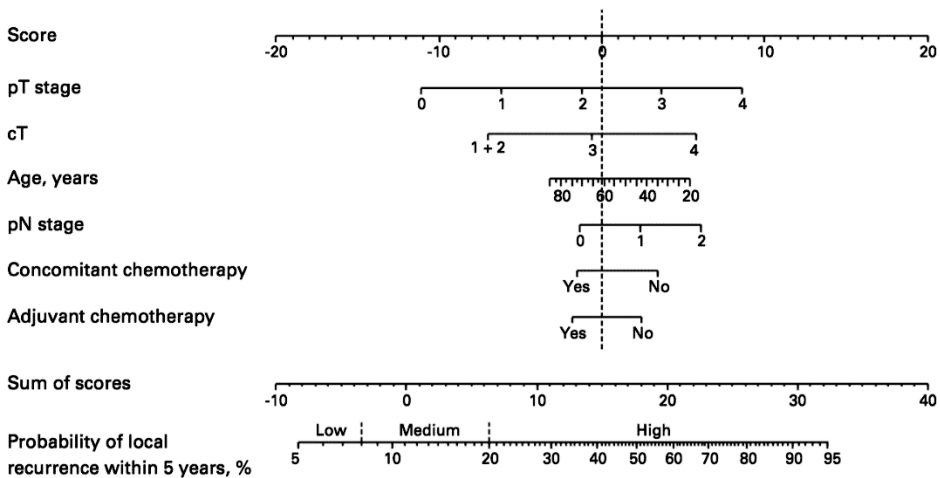
Figure 3 Example of a nomogram [8]

Decision support systems are neither intended nor suited as a replacement for the physician as a healthcare professional. They are designed to support the physician and the patient in making a more informed decision with respect to a particular treatment. The use of computer models to support healthcare professionals in their efforts is, of course, not new in radiation oncology. Physics-based computer models, with which doses can be better calculated than by hand, as well as radiobiology-based Normal Tissue Complications Probability (NTCP) and Tumour Control Probability (TCP) models to correlate the given dose with tumour control and toxicity, are commonplace within radiotherapy [27,28]. For example, geometrical models based on tumour volume alone have shown additional value next to classical TNM classification as well [29]. The new models emerging from Rapid Learning are a natural extension of this to patient outcomes. However, a key difference is that the Rapid Learning models

are more 'holistic' and multifactorial than the current physics- or radiobiology-based models, as they also take patient, tumour and non-radiotherapy factors into account [30]. For instance, a Rapid Learning model of radiation-induced oesophagitis shows that the risk for this toxicity not only depends on the dose to which the oesophagus is exposed, but also greatly increases if chemotherapy is given concomitantly [31]. Another example is that the survival of non-metastatic unresectable non-small cell lung cancer is better predicted by a multi-factorial model based on clinical and imaging variables, and even more when blood biomarkers are included [31,32]. In both cases the models outperform the prognostic value of TNM classification.

Evaluation

The underlying idea in Rapid Learning is that the application of knowledge acquired from routine data leads to predictability of treatment outcomes, meaning that these outcomes can be improved in terms of both effectiveness (achieving the desired result) and efficiency (the resources needed to achieve the result). Naturally, this needs to be continually evaluated, focusing on the question 'Is the outcome of the treatment as predicted?' Compared to the consensus- and evidence based guideline knowledge that is preferably constructed with (meta-analysis of) robust experimental data that are interpreted by multiple stakeholders including health care economists and patient representatives, the prediction models may suffer from confounders and selection bias. For Rapid Learning, having high-quality data with respect to outcomes is crucial. This implies the use of broadly accepted taxonomies such as RECIST or pathological Complete Response for tumour response [33], CTCAE for toxicity [34] and euroQoL for quality of life & utilities (which allow to calculate Quality Adjusted Life Year (QALY)) [35,36]. Naturally, keeping thorough records of treatment outcomes is important not only for Rapid Learning, but also for initiatives such as the quality registration system for lung cancer patients initiated by the Dutch Society for Radiotherapy and Oncology.

Discussion

Tailored cancer treatment is a necessity, to ensure not only that the individual patient receives the treatment that best suits his or her wishes, and to avoid under or overtreatment but also to optimise resources, so that the right resources

are being used for the right patients in healthcare in a broader sense. However, tailored cancer treatment is also a challenge: the great diversity of cancer patients and treatments implies that it is by no means always clear which choice leads to which treatment outcome. Especially in cases where the treatment options under consideration have no clear clinical advantage in the outcome, a shared decision-making process can be employed in order to make the most of patient preferences.

Tailored therapy is also necessary for radiotherapy. The radio-sensitivity of tumours and normal tissues is often unknown, certainly not homogeneous within an individual patient, and even less so between patients [37–40]. In addition, the range of treatment options and thus the number of decisions that need to be made within radiotherapy have risen sharply, largely due to technological innovations such as IMRT, VMAT, IGRT and particle therapy as well as innovative combinations with systemic and targeted treatments such as tyrosine inhibitors or monoclonal antibodies (e.g., Cetuximab). Opting for a particular radiation treatment on the basis of expected outcomes is therefore difficult, and the established guidelines and literature provide only limited support in this regard.

This article has discussed Rapid Learning as a means of support when deciding on a tailored radiation treatment. In essence, Rapid Learning involves reusing local, clinical, routine data to develop knowledge in the form of models that can predict treatment outcomes, and then clinically applying and carefully evaluating these models by way of Decision Support Systems. The hypothesis is that treatment outcomes obtained in the past can be used to predict future results.

Earlier attempts to introduce so-called 'expert systems' had mixed results. The proposed Rapid Learning methodology is different from the earlier attempts to deploy expert systems in several ways: it makes use of larger quantities of relevant data (e.g. the clinical patient population), as steadily more clinical data become available electronically in the clinical environment. This also enables validation in one's local practice which is a prerequisite for any expert system to be accepted, similar to commissioning and acceptance of treatment planning systems in radiotherapy. In contrast with expert systems, Rapid Learning employs quantitative models in addition to qualitative models. Finally, the de facto current expert system from "literature and guidelines based on clinical trials" has limited application to personalised medicine. This will drive the demand for more flexible and rapidly updated expert systems such as proposed in this review.

The Rapid Learning approach seems to contradict the principles of evidence-based medicine, in which treatment decisions are based solely on results obtained

from controlled clinical trials. In fact, it does not; both approaches are complementary (Fig. 4). This is compounded by the fact that Rapid Learning is based on results obtained from the less controlled setting of clinical practice. These different environments yield different insights. Controlled clinical trials primarily aim to identify small improvements in results between two treatments in a patient group that is as homogeneous as possible. In contrast, Rapid Learning will reveal major differences in treatment outcomes that stem from the heterogeneity of the patient group. It will be inferior in detecting minor differences in treatments due to the lower quality of the data recorded in clinical practice as compared to the same treatment in a clinical trial. In addition, Rapid Learning can be seen as an alternative for situations in which there are insufficient evidence to make decisions in line with the principles of evidence-based medicine. This is often the case with technological innovations; for instance, when considering the use of new techniques (e.g., IMRT, protons) in the field of radiotherapy [41].
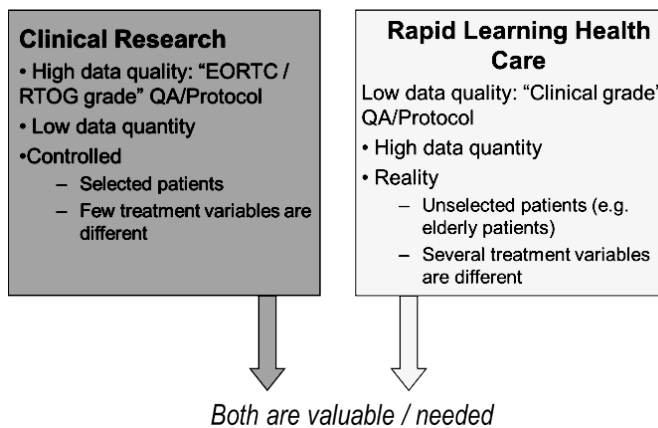
Figure 4 Complementary instead of contradictory approaches

Rapid Learning is new and still needs to prove its value as a supplement to traditional, evidence-based approaches. There are several developments that might help Rapid Learning change the way scientific evidence is viewed in medicine: (a) Technological advances will be created by larger and higher quality databases that link electronic health records with research databases, as well as the advent of the Semantic Web with increased interoperability and distributed learning approaches that enable learning from data without the need for data to leave the hospital; (b) The development by domain experts of qualitative criteria to evaluate evidence coming from large databases and rapid learning approach-

es; (c) The increased pressure and possible reimbursement from healthcare payers to use Decision Support Systems, especially for high cost treatments such as proton therapy; and (d) The development of "clinical grade" and certified commercial decision support systems.

Radiotherapy seems to be the ideal setting to study the value of Rapid Learning, given the field's high degree of computerisation, as well as its long use and acceptance of predictive models. Within clinical radiotherapy, models and planning systems should become available that make it possible to not only plan on the basis of physical dose and Dose Volume Histograms parameters, but also to explain the relationship with the expected clinical outcomes in individual patients. Translating knowledge to an individual patient is challenging, particularly in so-called preference-sensitive situations where there are trade-offs between options with more or less equally desirable outcomes, but in which different individuals may value differently e.g. in terms of side effects. As access to health-related information improves, patients have an increased desire to be in charge of their own life and health. Despite investment in efforts to improve the skills of clinicians, patients continue to report low levels of involvement [42]. There is indeed evidence level 1 from a Cochrane systematic review evaluating 86 studies involving 20,209 participants included in published randomised controlled trials demonstrating that decision aids increase people's involvement, support informed values-based choices in patient-practitioner communication, and improve knowledge and realistic perception of outcomes. We therefore believe the next step will be to integrate, whenever possible, Shared Decision Making approaches (see for example www.treatmentchoice.info; www.option grid.org) to include the patient perspective in the choice of best treatment [26].

Conflict of interest

We are not aware of any actual or potential conflicts of interest.

Acknowledgements

References

[1] Lambin P, Petit SF, Aerts HJ, van Elmpt WJ, Oberije CJ, Starmans MH, et al. From population to voxel-based radiotherapy: exploiting intra-tumour and intra-organ heterogeneity for advanced treatment of non-small cell lung cancer. Radiother Oncol 2009;2010:145–52.

[2] Koboldt DC, Fulton RS, McLellan MD. Comprehensive molecular portraits of human breast tumours. Nature 2012;490(7418):61–70.

[3] Seigneuric R, Starmans MH, Fung G, Krishnapuram B, Nuyten DS, van Erk A, et al. Impact of supervised gene signatures of early hypoxia on patient survival. Radiother Oncol 2007;83:374–82.

[4] Starmans MH, Chu KC, Haider S, Nguyen F, Seigneuric R, Magagnin MG, et al. The prognostic value of temporal in vitro and in vivo derived hypoxia gene-expression signatures in breast cancer. Radiother Oncol 2012;102:436–43.

[5] Starmans MH, Fung G, Steck H, Wouters BG, Lambin P, et al. A simple but highly effective approach to evaluate the prognostic performance of gene expression signatures. PLoS ONE 2011;6:e28320.

[6] Starmans MH, Lieuwes NG, Span PN, Haider S, Dubois L, Nguyen F, et al. Independent and functional validation of a multi-tumour-type proliferation signature. Br. J. Cancer 2012;107:508–15.

[7] Starmans MH, Zips D, Wouters BG, Baumann M, Lambin P. The use of a comprehensive tumour xenograft dataset to validate gene signatures relevant for radiation response. Radiother Oncol 2009;92:417–22.

[8] Voets AM, Oberije C, Struijk RB, Reymen B, De Ruyck K, Thierens H, et al. No association between TGF-beta1 polymorphisms and radiation-induced lung toxicity in a European cohort of lung cancer patients. Radiother Oncol 2012;105:296–8.

[9] Movsas B, Moughan J, Owen J, Coia LR, Zelefsky MJ, Hanks G, et al. Who enrolls onto clinical oncology trials? A radiation patterns of care study analysis. Int J Radiat Oncol Biol Phys 2007;68:1145–50.

[10] Grand MM, O'Brien PC. Obstacles to participation in randomised cancer clinical trials: a systematic review of the literature. J Med Imag Radiat Oncol 2012;56:31–9.

[11] Murthy VH, Krumholz HM, Gross CP. Participation in cancer clinical trials: race-, sex-, and age-based disparities. JAMA 2004;291:2720–6.

[12] Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? PLoS Med 2010;7:e1000326.

[13] Abernethy AP, Etheredge LM, Ganz PA, Wallace P, German RR, Neti C, et al. Rapid-learning system for cancer care. J Clin Oncol 2010;28:4268–74.

[14] Dehing-Oberije C, Yu S, De Ruysscher D, Meersschout S, Van Beek K, Lievens Y, et al. Development and external validation of prognostic model for 2-year survival of non-small-cell lung cancer patients treated with chemoradiotherapy. Int J Radiat Oncol Biol Phys 2009;74:355–62.

[15] Egelmeer AG, Velazquez ER, de Jong JM, Oberije C, Geussens Y, Nuyts S, et al. Development and validation of a nomogram for prediction of survival and local control in laryngeal carcinoma patients treated with radiotherapy alone: a cohort study based on 994 patients. Radiother Oncol 2011;100:108–15.

[16] Jimenez MF, van Baardwijk A, Aerts HJ, De Ruysscher D, Novoa NM, Varela G, et al. Effectiveness of surgery and individualized high-dose hyperfractionated accelerated radiotherapy on survival in clinical stage I non-small cell lung cancer. A propensity score matched analysis. Radiother Oncol 2010;97:413–7.

[17] Roelofs E, Engelsman M, Rasch C, Persoon L, Qamhiyeh S, de Ruysscher D, et al. Results of a multicentric in silico clinical trial (ROCOCO): comparing radiotherapy with photons and protons for non-small cell lung cancer. J Oncol 2012;7:165–76.

[18] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441–6.

[19] Roelofs E, Persoon L, Nijsten S, Wiessler W, Dekker A, Lambin P. Benefits of a clinical data warehouse with data mining tools to collect data for a radiotherapy trial. Radiother Oncol 2013. doi:pii: S0167-8140(12)00444-6.

[20] van Elmpt W, Nijsten S, Mijnheer B, Dekker A, Lambin P. The next step in patient-specific QA: 3D dose verification of conformal and intensity-modulated RT based on EPID dosimetry and Monte Carlo dose calculations. Radiother Oncol 2008;86:86–92.

[21] Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128–38.

[22] Valentini V, van Stiphout RG, Lammering G, Gambacorta MA, Barba MC, Bebenek M, et al. Nomograms for predicting local recurrence, distant metastases, and overall survival for patients with locally advanced rectal cancer on the basis of European randomized clinical trials. J Clin Oncol 2011;29:3163–72.

[23] De Ruysscher D, Dehing C, Yu S, Wanders R, Ollers M, Dingemans AM, et al. Dyspnea evolution after high-dose radiotherapy in patients with non-small cell lung cancer. Radiother Oncol 2009;91:353–9.

[24] De Ruysscher D, Houben A, Aerts HJ, Dehing C, Wanders R, Ollers M, et al. Increased (18)F-deoxyglucose uptake in the lung during the first weeks of radiotherapy is correlated with subsequent radiation-induced lung toxicity (RILT): a prospective pilot study. Radiother Oncol 2009;91:415–20.

[25] Dehing-Oberije C, De Ruysscher D, van Baardwijk A, Yu S, Rao B, Lambin P, et al. The importance of patient characteristics for the prediction of radiation-induced lung toxicity. Radiother Oncol 2009;91:421–6.

[26] Stacey D, Bennett CL, Barry MJ, Col NF, Eden KB, Holmes-Rovner M, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2011:CD001431.

[27] Bentzen SM, Dorr W, Gahbauer R, Howell RW, Joiner MC, Jones B, et al. Bioeffect modeling and equieffective dose concepts in radiation oncology– terminology, quantities and units. Radiother Oncol 2012;105:266–8.

[28] De Leeuw AA, Van de Kamer JB, Moerland MA, Philippens ME, Jurgenliemk-Schulz IM. The effect of alternative biological modelling parameters (alpha/ beta and half time of repair T(1/2)) on reported EQD2 values in the treatment of advanced cervical cancer. Radiother Oncol 2011;101:337–42.

[29] Guo R, Sun Y, Yu XL, et al. Is primary tumor volume still a prognostic factor in intensity modulated radiation therapy for nasopharyngeal carcinoma? Radiother Oncol 2012;104:294–9.

[30] Lambin P, van Stiphout RG, Starmans MH, Rios-Velazquez E, Nalbantov G, Aerts HJ, et al. Predicting outcomes in radiation oncology – multifactorial decision support systems. Nature Rev Clin Oncol 2013;10:27–40.

[31] Dehing-Oberije C, De Ruysscher D, Petit S, Van Meerbeeck J, Vandecasteele K, De Neve W, et al. Development, external validation and clinical usefulness of a practical prediction model for radiation-induced dysphagia in lung cancer patients. Radiother Oncol 2010;97:455–61.

[32] Dehing-Oberije C, Aerts H, Yu S, De Ruysscher D, Menheere P, Hilvo M, et al. Development and validation of a prognostic model using blood biomarker information for prediction of survival of non-small-cell lung cancer patients treated with combined chemotherapy and radiation or radiotherapy alone (NCT00181519, NCT00573040, and NCT00572325). Int J Radiat Oncol Biol Phys 2011;81:360–8.

[33] Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45:228–47.

[34] Trotti A, Colevas AD, Setser A, Rusch V, Jaques D, Budach V, et al. CTCAE v3.0: development of a comprehensive grading system for the adverse effects of cancer treatment. Semin Radiat Oncol 2003;13:176–81.

[35] Grutters JP, Kessels AG, Pijls-Johannesma M, De Ruysscher D, Joore MA, Lambin P. Comparison of the effectiveness of radiotherapy with photons, protons and carbon-ions for non-small cell lung cancer: a meta-analysis. Radiother Oncol 2010;95:32–40.

[36] Peeters A, Grutters JP, Pijls-Johannesma M, Reimoser S, De Ruysscher D, Severens JL, et al. How costly is particle therapy? Cost analysis of external beam radiotherapy with carbon-ions, protons and photons. Radiother Oncol 2010;95:45–53.

[37] Aerts HJ, Bussink J, Oyen WJ, van Elmpt W, Folgering AM, Emans D, et al. Identification of residual metabolic-active areas within NSCLC tumours using a pre-radiotherapy FDG-PET-CT scan: a prospective validation. Lung Cancer 2012;75:73–6.

[38] van Loon J, Janssen MH, Ollers M, Aerts HJ, Dubois L, Hochstenbag M, et al. PET imaging of hypoxia using [18F]HX4: a phase I trial. Eur J Nucl Med Mol Imag 2010;37:1663–8.

[39] Mortensen LS, Johansen J, Kallehauge J, Primdahl H, Busk M, Lassen P, et al. FAZA PET/CT hypoxia imaging in patients with squamous cell carcinoma of the head and neck treated with radiotherapy: results from the DAHANCA 24 trial. Radiother Oncol 2012;105:14–20.

[40] Zips D, Zophel K, Abolmaali N, Perrin R, Abramyuk A, Haase R, et al. Exploratory prospective trial of hypoxia-specific PET imaging during radiochemotherapy in patients with locally advanced head-and-neck cancer. Radiother Oncol 2012;105:21–8.

[41] Langendijk JA, Lambin P, De Ruysscher D, Widder J, Bos M, Verheij M. Selection of patients for radiotherapy with protons aiming at reduction of side effects: the model-based approach. Radiother Oncol 2013.

[42] Stiggelbout AM, Van der Weijden T, De Wit MP, Frosch D, Legare F, Montori VM, et al. Shared decision making: really putting patients at the centre of healthcare. BMJ 2012;344:e256.

[43] Deasy JO, Bentzen SM, Jackson A, Ten Haken RK, Yorke ED, Constine LS, et al. Improving normal tissue complication probability models: the need to adopt a "data-pooling" culture. Int J Radiat Oncol Biol Phys. 2010;76(3 Suppl):S151–4.

Chapter 3

CLINICAL DATA WAREHOUSING

*Benefits of a clinical data warehouse with data mining
tools to collect data for a radiotherapy trial*

Erik Roelofs*, Lucas Persoon*, Sebastiaan Nijsten, Wolfgang Wiessler,
André Dekker*, Philippe Lambin*; *Equal contribution

Abstract

Introduction: Collecting trial data in a medical environment is at present mostly performed manually and therefore time-consuming, prone to errors and often incomplete with the complex data considered. Faster and more accurate methods are needed to improve the data quality and to shorten data collection times where information is often scattered over multiple data sources. The purpose of this study is to investigate the possible benefit of modern data warehouse technology in the radiation oncology field.

Material and methods: In this study, a Computer Aided Theragnostics (CAT) data warehouse combined with automated tools for feature extraction was benchmarked against the regular manual data-collection processes. Two sets of clinical parameters were compiled for non-small cell lung cancer (NSCLC) and rectal cancer, using 27 patients per disease. Data collection times and inconsistencies were compared between the manual and the automated extraction method.

Results: The average time per case to collect the NSCLC data manually was 10.4 ± 2.1 min. and 4.3 ± 1.1 min. when using the automated method ($p<0.001$). For rectal cancer, these times were 13.5 ± 4.1 and 6.8 ± 2.4 min., respectively ($p<0.001$). In 3.2% of the data collected for NSCLC and 5.3% for rectal cancer, there was a discrepancy between the manual and automated method.

Conclusions: Aggregating multiple data sources in a data warehouse combined with tools for extraction of relevant parameters is beneficial for data collection times and offers the ability to improve data quality. The initial investments in digitizing the data are expected to be compensated due to the flexibility of the data analysis. Furthermore, successive investigations can easily select trial candidates and extract new parameters from the existing databases.

Introduction

Collecting data in a medical environment for research purposes is time consuming, prone to errors and often incomplete when complex data, such as dosimetric data is concerned [1,2]. Faster and more accurate access to medical information is required to improve the data quality, shorten data-collection times and reduce cost. Furthermore, conducting trials in multiple centres and data-sharing is required in numerous research projects to validate trial results, enable the collection of larger datasets of trial patients with a rare disease or to conduct *in silico* trials [1,3-6]. Although the medical community has undertaken numerous efforts to digitize its patient and treatment documentation, data are still predominantly collected from paper charts. And when data are collected digitally, they are in most cases unstructured and distributed over multiple data sources.

With the growing number of diagnostic and therapeutic modalities comes an increasing demand for clinical trials to provide the evidence base for these new options and to provide guidance to healthcare providers. However, with only around 3% of the adult cancer patients included in clinical trials this is a very hard task [7-9]. The use of healthcare information systems based on multi-parametric electronic medical record databases (EMR) and data mining tools will greatly enhance clinical research in oncology and more specifically in radiation oncology and facilitate trials with easy patient selection and improved data quality.

Because data is often scattered and unstructured throughout a medical care organization, data warehouse technology is suitable to combine data sources, validate consistency and share data with other researchers [10-12]. It can integrate various information systems in the healthcare enterprise and offers the ability to have specific data structures/views for different investigations, which may deviate from the structure of the operational systems.

Benefits of data warehousing have been described on multiple levels, such as time-saving for users, improved quantity and quality of information, informed decision-making, improvement of business processes and ultimately support for the accomplishment of strategic business objectives [13]. Data warehousing is increasingly used in healthcare to provide the tools for decision making and individualizing disease management [14,15]. Furthermore, it is essential to facilitate the (translational) research that is needed to develop new treatment programs and support clinical trials (i.e. with quality assurance programmes) [16-20].

To this end, we developed a research portal with an industrial partner to integrate the essential medical data sources and offer automated data extraction

tools for research purposes. In this study, we test the performance of this portal with data mining tools against the manual collection process for clinical trials. Performance is measured in time expenses and data quality to target the hypothesis that these will decrease and improve, respectively, by the use of a data warehouse.

Methods and Materials

The Computer Aided Theragnostics (CAT) research portal

Together with Siemens Knowledge Solutions (Malvern, PA, USA) we developed a Computer Aided Theragnostics (CAT) research portal. It extracts medical data from the connected systems via a synchronization manager (sync manager) and stores the data centrally in a data warehouse. The operational, patient-centric structure is converted into a disease-centric structure suitable for research. In our radiotherapy department, the sync manager extracts data from various sources: 1) the electronic medical record (EMR), which is either a structured or unstructured database, 2) the RT picture archiving communication system (PACS), consisting of diagnostic imaging and treatment DICOM RT data such as the treatment plans (RTPLAN), predicted 3D dose matrices (3D-RTDOSE), delineations (RTSTRUCT) and digitally reconstructed radiographs used for setup verification (RTIMAGE), and 3) the Record and Verify system (R&V) containing the actual delivered treatment parameters.

The CAT research portal currently has four core user functionalities (Figure 1). The first module is the query builder, which is a tool for the visual creation and execution of queries as well as viewing the query results and exporting them in XML format. An integrated data-browser (second module) is available where individual cases can be reviewed. Thirdly, the system offers an electronic case report form (eCRF) module. This enables researchers to use the system for clinical trials. The module is built in such a way that case report forms can be pre-populated with data extracted from the clinical databases acquired during the sync procedure. The fourth module is a private data-store for researchers to upload parameters of interest not collected by either eCRFs or clinical systems.

In addition to the EMR, PACS and R&V systems, our department designed custom-made, fully automated workflows in a framework called DIGITrans (www.mistir.info/digitrans). DIGITrans offers a user interface and background services for data generation, extraction and transportation for both clinical and

research purposes. During daily clinical practice, 2D dose-guided radiotherapy workflows are present [21], while for research users, workflows are defined to extract parameters from for instance dose-volume histograms (DVH) and store the results as DICOM-objects in the PACS system. These workflows are DICOM-driven and can convert, validate and transfer DICOM objects throughout the entire department and can easily be adapted to support new parameters or modalities. Parameters of interest are stored in a structured database. Currently, DVH parameters such as mean lung dose (MLD), structure volumes and different volume and dose parameters are stored in a separate structured database. A similar extraction is done from PET scans to derive and store standardized uptake values (SUV).
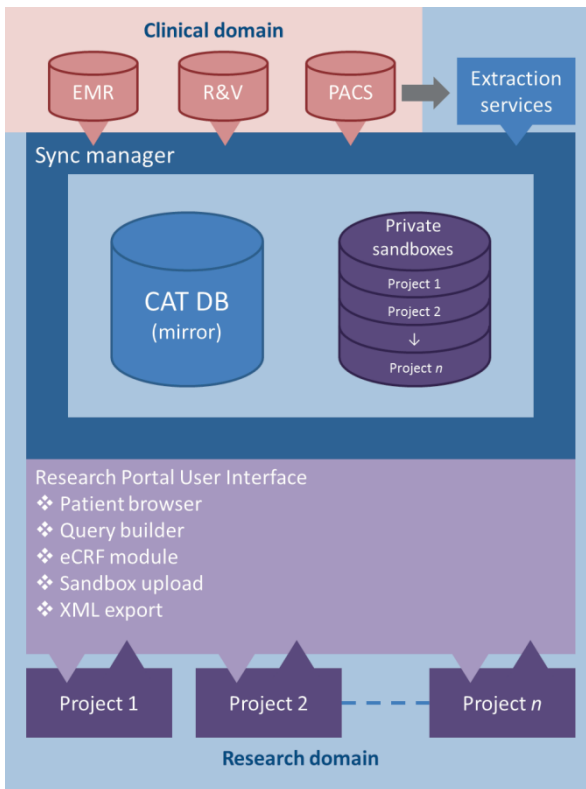
Figure 1: Schematic overview of the CAT data warehouse / research portal. The system synchronizes data from clinical data sources and custom services. It is also capable of collecting data for trials and data collected for other research purposes. For data export, several modules exist in the system and are easily accessible by web-technology (i.e. the patient browser, query builder and an electronic case report form XML export).

The current data management process

The current trial data management process in our institute is largely manual and when data is acquired electronically still the case report forms have to be filled in manually. Currently, a list of patients and the parameters that need to be collected is supplied to the data manager, who will then identify from which system or document the data needs to be extracted. Next, the data is collected manually from each system or chart. The data is noted in an electronic document such as Microsoft Excel, Access or in an electronic case report form (eCRF). Data is collected using the Good Clinical Practice (GCP) guidelines as a reference [22,23].

Collection of some of the parameters mentioned requires domain and application specialists. For these parameters, the data managers do not collect data (e.g. DVH or SUV parameters, which require recalculation), but the collection is done separately by radiation technologists.

Experimental setup

Two hypothetical trials with representative sets of parameters were defined for non-small cell lung cancer (NSCLC) and rectal cancer (Table 1). Next, a comparison was made between the manual data collection process and the automatic data warehouse based method. The datasets were heterogeneous so that data had to be collected from the paper charts (manually) or EMR (automatically), the R&V system and XiO/TrueD (manually) or the PACS. For the measurements, the parameter sets were split into two groups; 1) the "lookup" group (chart/EMR and R&V) and 2) the "recalculation" group (XiO/TrueD and PACS).

The manual recalculation of DVH parameters was conducted with the treatment planning system (TPS) XiO (CMS Software, Elekta, XiO 4.34.02, St. Louis, MO), using the plans available on the TPS system. For the manual recalculation of SUV values we used the commercially available TrueD software (TrueD VC50, Siemens Medical, Erlangen, Germany). The PET images were re-imported and volumes of interest (VOIs) were created again to retrieve the SUV data.

For the automatic recalculated parameters, a query was defined once using the CAT research portal and the result was exported in XML format. From this query result, the original CT/PET images, structures and dose were sent from the PACS to the DIGITrans workflows to automatically extract the DVH and SUV data.

Table 1: Parameters collected for the NSCLC and rectal cancer groups. The last columns show which data was looked up where and from which source the data was recalculated.

| Parameter	NSCLC	Rectum	Source		Action
			Manual	Automatic	
Gender	√	√	Chart	EMR	Looked up
WHO score	√	√	Chart	EMR	
TNM staging	√	√	Chart	EMR	
Chemo therapy	√	√	Chart	EMR	
Nr. positive lymph nodes	√	√	Chart	EMR	
Tumour PA	√	√	Chart	EMR	
pCR		√	Chart	EMR	
Survival	√	√	Chart	EMR	
Total delivered dose	√	√	R&V	R&V	
Overall treatment time	√	√	R&V	R&V	
GTV volume	√	√	XiO	PACS	Recalculated
V_5	Lungs[1]		XiO	PACS	
V_{20}	Lungs		XiO	PACS	
V_{40}		Bladder	XiO	PACS	
MLD	√[2]		XiO	PACS	
SUV Max		Tumour	TrueD	PACS	
SUV Mean		Tumour	TrueD	PACS	

[1] V_5 and V_{20} data for the lungs were calculated with both lungs minus the PTV

[2] MLD data for the lungs were calculated with both lungs minus the GTV

Furthermore, data quality was evaluated for all experiments by scoring discrepancies between both the CAT and current manual process. To decide which method was correct, each deviation was looked up again in its reference system or validated by another observer.

The primary end-point of the experiment was data collection time. For the sample size calculation, we hypothesized a difference of 3 minutes with a standard deviation of 3 minutes using a paired t-test. We wanted to achieve a power of 0.99 and an α=0.01. Based on these assumptions, the sample size was determined at n=27. The secondary end-point of data quality was evaluated by comparing the output from the CAT method against the first human observer. The output of the human observer was validated by a second human observer and inconsistencies were resolved by a third look-up. The third look-up was not taken into account in the data collection time measurements. After these comparisons between observers, a GCP compliant collected data-set was acquired and this set was also compared against the CAT method.

Results

In Figure 2, the average data collection times for NSCLC (a) and rectum cancer (b) are shown for the manual versus the CAT groups. The figure shows the distinction between the data that was looked up (in the EMR and R&V) versus recalculated (from PACS/TPS data). On average, the total collection time for the manual method for the individual NSCLC cases was 10.4±2.1 min. while for the CAT method this was 4.3±1.1 min. (p<0.001). For rectal cancer, this was 13.5±4.1 min. for the manual collection and 6.8±2.4 min. for the CAT method (p<0.001) (Table 2).
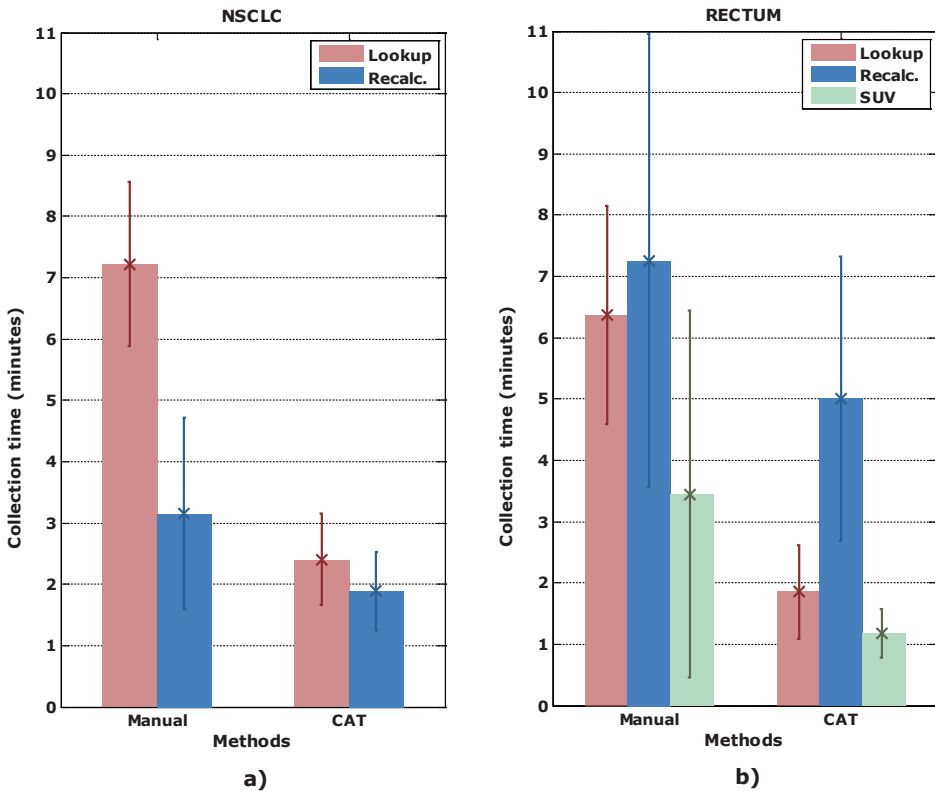
Figure 2: Average manual versus CAT collection times (in min.) for the a) NSCLC and b) Rectum cases. The parameters that were looked up in the EMR and R&V system are displayed in pink and labelled "Lookup". In blue (labelled "Recalc."), the parameters are shown that were recalculated. The error bars show the standard deviations. For the rectum cases, the collection times for SUV data only show the large variability in the contribution to the recalculated parameters (in green and labelled "SUV").

Table 2: Comparison of the total data collection times (mean±SD in min.) for the manual versus the CAT method.

Tumour type	Manual (min)	CAT (min)	p-value
NSCLC	10.4±2.1	4.3±1.2	<0.001
Rectum	13.5±4.1	6.9±2.3	<0.001

The main difference between the manual and automatic collection times for both the NSCLC and rectum cases were caused by the lookup parameters. The data warehouse offered one central interface where all parameters were present, while the manual method required the data managers to gather and analyse the paper charts.

For the NSCLC cases there was a difference between manual and automated collection times of the DVH parameters. This was caused by the fact that the V_5 and V_{20} are not used in daily clinical practice and their DVHs needed to be re-calculated in the TPS.

For the rectal cancer cases there was a difference compared to the NSCLC cases with respect to the collection time of recalculated parameters. This difference was mainly caused by the extraction of the SUV_{mean} and SUV_{max}. In the manual method, VOIs needed to be re-drawn for all cases because the used system did not support import from structures or back-upped data. As can be seen in Figure 2b (SUV data), the standard deviation is very large (±3.0 min). This is due to the effect that in some cases the auto-segmentation delineated neighbouring organs such as the bladder and contours needed to be edited manually. To determine the data-quality, we compared the SUV parameters with the newly generated structure-set.

There was no difference found in the extraction of DVH parameters between both methods, however, the extraction took significantly longer when compared to the NSCLC extraction. This was because in daily clinical practice the bladder is normally not delineated and therefore not available for both methods and was delineated in the TPS. The time needed for the bladder delineation was taken into account for both methods and therefore there is no difference between both methods for this extraction.

Manual data collection was performed according to GCP guidelines. This means that data is collected by two observers and then validated. After validation between observers, we found that in the NSCLC set there was a mismatch between manually collected data and automatic in 3.2% of all different parameters collected. The mismatch before validation for NSCLC was 10.3%. Using the GCP collection process thus resolved 69% of the initial mismatches.

We found a particular bad concordance between observers for the manual collection of the overall treatment time in days. This was due to the fact that some observers erroneously recalculated the data while others looked it up in the R&V system. This interpretation difference explained most of the drop of mismatching data after validation. The remaining differences could be explained primarily by a mismatch between DVH parameters. We found that in most cases the DVH parameters that mismatched before validation were based on a different treatment plan than actually used for treatment or were calculated on a wrong volume. The V_5 and V_{20} were defined using $Lung_{Left}$ – $Lung_{Right}$ – $PTV(s)$ volumes while the MLD had to be calculated on $Lung_{Left}$ – $Lung_{Right}$ – $GTV(s)$ volumes. Before validation, the V_5 and V_{20} were in some cases extracted from the DVH of where the MLD is calculated on. After the validation between observers, the remaining 3.2% mismatching parameters indicate the true mismatch between both methods. We found that in more than 90% of these cases the automatic method retrieved the correct data.

For the rectal cancer set we observed similar results. We observed data differences in 8.8% of all data collected. Again, we saw a bad concordance of the overall treatment time between the manual and automatic method. The other deviations were mainly found in the recalculation of SUV parameters. The cause for the deviation could be found in the implementation differences between the automated mining tool and the PET analysis software.

In the PET analysis software, a threshold is calculated by taking into account the signal-to-background ratio calculated as Threshold = $0.7813*(SUV_{maxTumour}/SUV_{meanBackground})^{-0.299}$. This threshold is then used to draw the contour on the PET and all the voxels with an SUV above the threshold are used to calculate the SUV_{mean}. The automated method however uses the exported contour to calculate the mean SUV from all the voxels within the exported contour. Due to interpolation and inaccuracy of the contour the SUV_{mean} of the automated method was lower than with the analysis software. This problem was compensated to search the contours for the background contours, calculate the threshold and use the threshold in the same way as the PET analysis tool. This explained 31% of the mismatches.

The remaining data inconsistencies were either human error(s), for example in one case we found a different total delivered dose, or were caused by erroneous DICOM objects.

Discussion

Our results show that using a data-warehouse shortens data-collection times significantly and can be of help to improve data quality because data in the CAT data warehouse is captured from all clinical reference systems and validated before storage in the warehouse. Especially the use of the in-house developed mining tools support easy extraction or recalculation of parameters of interest (e.g. V_{20}, V_5, SUV_{mean} and SUV_{max}) and eliminates manual steps. However, as the results show with the extraction of the SUV parameters great care must be taken when using these kinds of tools in an automated way. It is very important to carefully validate and approve them before extensive research use.

The strength of the CAT data warehouse is that it can combine, validate and present data from distributed databases in a uniform way. Furthermore, the system is designed to have an external interface to share data among other institutes. Data is easily accessible for researchers via a web-interface with several extra options such as a patient browser and a query builder. Occasionally, expert knowledge is necessary to maintain the system, because of interoperability issues due to updates of source systems.

When looking at the manual data collection process we have found that there was a large inter-observer variability. This explains most of the data inconsistencies between the automated and manual collection methods. Strict guidelines such as GCP [22,23] are of great importance to ensure data quality. We found that the collection by two observers and the monitoring and validation is vital for data quality because it resolves conflicts between observers. From previous studies [24,25] it can be concluded that, by using intelligent free-text searching/mining and machine learning techniques approaches to retrieve the data in a data warehouse, data can automatically be validated and true data values can be identified. Retrieving this data manually can introduce inconsistencies or missing data, which we have seen in some cases in this study. This is confirmed by Prokosch et al. [16] who mention in their second challenge that electronic data capture for trials provides significant benefits over manual collection. One other advantage is the possibility to include real-time patient outcome data, for instance by directly linking to a national registry or including properly capture patient-reported outcome [26,27].

Using the research portal for trial data collection with eCRFs using extracted data from local databases, could possibly be a better start of the data collection and even replace the first observer. The electronically captured data could then

immediately be validated by the monitor. By using this approach the data collection time can be minimized. Although pre-population of data can induce data capturing problems as mentioned by Kush et al. [2] this can be solved by the requirement that the responsible eCRF reporter verifies all pre-populated data. This can be embedded in the user interface workflow. New technologies such as developed by Rao et al. [24] can be used to validate data at the synchronization process delivering high quality data to the user who only needs to validate the data for the clinical trial.

The CAT research portal is an institute-specific data warehouse that can also be used to share data in a privacy-preserving and semantic-interoperable manner using internationally accepted data exchange standards and ontologies such as described in an accompanying study [28,29], addressing the growing need for standardized data exchange between medical centres [1,3,12]. Richesson et al. [30] describe that there are currently overlaps, challenges and gaps in the standards developed for research and clinical purposes for data retrieval at a local scale and data sharing at a global scale. They conclude that data should be shared between research and clinics where possible but also that the clinical data will have gaps and that the importance of data standards for clinical research is underestimated. In one of a series of vision papers by the QUANTEC group, Deasy et al. [31] suggest that data-sharing and -pooling can raise the quantity as well as the quality of clinical data for data mining purposes. We strongly share this believe.

A data warehouse probably reduces cost for the research organization. As can be seen from the results from this study, data collection time can be reduced significantly. Although the fact that clinical trials usually run for several years the time spent for data collection can be reduced by more than 50%, which should translate into a costs reduction of data management. However, to set up and maintain systems like this will have financial implications. The initial installation is the most expensive part but can be spread over several projects in multiple years. The operational cost of a system is dependent on the organizations needs and is variable. For example, for typical projects datasets of several hundreds of patients are needed [32-35]. It was estimated in one of these studies that approximately 45 hours of data management can be saved, based on a dataset of 400 patients, which is 6-7 min. per case. This is in concordance with our findings.

For the manual collection process, we measured the time to collect the actual data when the source data was available in the sub-systems (i.e. the treatment plan was available in the TPS). However, in retrospective studies this might not

be the case and data might have to be restored from back-ups or archives. This additional time has not been taken into account in this study because of the fact that the retrieval of data differs heavily for various backup-and-restore systems. Most often PACS data resides on disks, while the TPS data is archived on tape or slow disks. Therefore, including restoration of data would likely increase manual recalculation times significantly.

A further benefit of data warehousing is to be expected when considering research and trials in the field of Radiomics [36], for instance. This involves the high-throughput extraction of large amounts of image features from clinical images. These quantitative imaging features are increasingly used in treatment planning and for monitoring patient outcome [37,38]. Because this is a rapidly evolving research area, new features are added frequently. Therefore, a data warehouse with easily modified data extraction services is likely to be far more beneficial when compared to manual data collection.

Conclusions

Information technology solutions such as the CAT data warehouse improve clinical research for radiotherapy by reducing the time needed to collect necessary data and by improving the quality of the data collected.

Acknowledgements

References

[1] Klein A, Prokosch HU, Muller M, Ganslandt T. Experiences with an interoperable data acquisition platform for multi-centric research networks based on HL7 CDA. Methods Inf Med 2007;46:580-585.

[2] Kush R, Alschuler L, Ruggeri R, et al. Implementing Single Source: the STARBRITE proof-of-concept study. J Am Med Inform Assoc 2007;14:662-673.

[3] Knaup P, Garde S, Merzweiler A, et al. Towards shared patient records: an architecture for using routine data for nationwide research. Int J Med Inform 2006;75:191-200.

[4] Roelofs E, Persoon L, Qamhiyeh S, et al. Design of and technical challenges involved in a framework for multicentric radiotherapy treatment planning studies. Radiother Oncol 2010;97:567-571.

[5] Bosmans G, Buijsen J, Dekker A, et al. An "in silico" clinical trial comparing free breathing, slow and respiration correlated computed tomography in lung cancer patients. Radiother Oncol 2006;81:73-80.

[6] Ollers M, Bosmans G, van Baardwijk A, et al. The integration of PET-CT scans from different hospitals into radiotherapy treatment planning. Radiother Oncol 2008;87:142-146.

[7] Movsas B, Moughan J, Owen J, et al. Who enrolls onto clinical oncology trials? A radiation Patterns Of Care Study analysis. Int J Radiat Oncol Biol Phys 2007;68:1145-1150.

[8] Murthy VH, Krumholz HM, Gross CP. Participation in cancer clinical trials: race-, sex-, and age-based disparities. JAMA 2004;291:2720-2726.

[9] Grand MM, O'Brien PC. Obstacles to participation in randomised cancer clinical trials: a systematic review of the literature. J Med Imaging Radiat Oncol 2012;56:31-39.

[10] Wisniewski MF, Kieszkowski P, Zagorski BM, Trick WE, Sommers M, Weinstein RA. Development of a clinical data warehouse for hospital infection control. J Am Med Inform Assoc 2003;10:454-462.

[11] Weber DC, Poortmans PMP, Hurkmans CW, Aird E, Gulyban A, Fairchild A. Quality assurance for prospective EORTC radiation oncology trials: The challenges of advanced technology in a multicenter international setting. Radiother Oncol 2011;100:150-156.

[12] El Fadly A, Rance B, Lucas N, et al. Integrating clinical research with the Healthcare Enterprise: from the RE-USE project to the EHR4CR platform. J Biomed Inform 2011;44 Suppl 1:S94-102.

[13] Watson HJ, Goodhue DL, Wixom BH. The benefits of data warehousing: why some organizations realize exceptional payoffs. Information & Management 2002;39:491-502.

[14] Schubart JR, Einbinder JS. Evaluation of a data warehouse in an academic health sciences center. Int J Med Inform 2000;60:319-333.

[15] Rubin DL, Desser TS. A data warehouse for integrating radiologic and pathologic data. J Am Coll Radiol 2008;5:210-217.

[16] Prokosch HU, Ganslandt T. Perspectives for medical informatics. Reusing the electronic medical record for clinical research. Methods Inf Med 2009;48:38-44.

[17] Sarkar IN. Biomedical informatics and translational medicine. J Transl Med 2010;8:22.

[18] Gaze MN, Boterberg T, Dieckmann K, et al. Development of an electronic database for quality assurance of radiotherapy in the International Society of Paediatric Oncology (Europe) high risk neuroblastoma study. Radiother Oncol 2010;97:593-595.

[19] Fairchild A, Aird E, Fenton PA, et al. EORTC Radiation Oncology Group quality assurance platform: Establishment of a digital central review facility. Radiother Oncol 2012;103:279-286.

[20] Bekelman JE, Deye JA, Vikram B, et al. Redesigning radiotherapy quality assurance: opportunities to develop an efficient, evidence-based system to support clinical trials-report of the national cancer institute work group on radiotherapy quality assurance. Int J Radiat Oncol Biol Phys 2012;83:782-790.

[21] Nijsten SM, van Elmpt WJ, Mijnheer BJ, et al. Prediction of DVH parameter changes due to setup errors for breast cancer treatment based on 2D portal dosimetry. Med Phys 2009;36:83-94.

[22] Mathieu M. Good Clinical Practice: A Question & Answer Reference Guide: Parexel Intl Corp, 2009.

[23] Pieterse H. Richtsnoer voor Good Clinical Practice (CPMP/ICH/95). Heerhugowaard: Profess Medical Consultancy B.V., 2010.

[24] Rao BR, Sandilya S, Niculescu R, Germond C, Goel A. Mining time-dependent patient outcomes from hospital patient records. Proc AMIA Symp 2002:632-636.

[25]   Branson A, Hauer T, McClatchey R, Rogulin D, Shamdasani J. A data model for integrating heterogene-
       ous medical data in the Health-e-Child project. Stud Health Technol Inform 2008;138:13-23.

[26]   Wong K, Huang SH, O'Sullivan B, et al. Point-of-care outcome assessment in the cancer clinic: audit of
       data quality. Radiother Oncol 2010;95:339-343.

[27]   Ho KF, Farnell DJ, Routledge JA, et al. Comparison of patient-reported late treatment toxicity (LENT-
       SOMA) with quality of life (EORTC QLQ-C30 and QLQ-H&N35) assessment after head and neck radio-
       therapy. Radiother Oncol 2010;97:270-275.

[28]   Dekker A, Roelofs E, Meldolesi E, Valentini V, Lambin P. Ad-hoc data sharing infrastructure for radio-
       therapy research collaboration: A tool for multicentric clinical research. Radiother Oncol 2011;99:S155-
       S156.

[29]   Roelofs E, Dekker A, Meldolesi E, van Stiphout RGPM, Valentini V, Lambin P. International data-
       sharing for radiotherapy research: an open-source based infrastructure for multicentric clinical data
       mining. Radiother Oncol 2014;110(2):370–374.

[30]   Richesson RL, Krischer J. Data standards in clinical research: gaps, overlaps, challenges and future direc-
       tions. J Am Med Inform Assoc 2007;14:687-696.

[31]   Deasy JO, Bentzen SM, Jackson A, et al. Improving normal tissue complication probability models: the
       need to adopt a "data-pooling" culture. Int J Radiat Oncol Biol Phys 2010;76:S151-154.

[32]   Jayasurya K, Fung G, Yu S, et al. Comparison of Bayesian network and support vector machine models
       for two-year survival prediction in lung cancer patients treated with radiotherapy. Med Phys
       2010;37:1401-1407.

[33]   Dehing-Oberije C, De Ruysscher D, van der Weide H, et al. Tumor volume combined with number of
       positive lymph node stations is a more important prognostic factor than TNM stage for survival of non-
       small-cell lung cancer patients treated with (chemo)radiotherapy. Int J Radiat Oncol Biol Phys
       2008;70:1039-1044.

[34]   Dehing-Oberije C, Yu S, De Ruysscher D, et al. Development and external validation of prognostic
       model for 2-year survival of non-small-cell lung cancer patients treated with chemoradiotherapy. Int J
       Radiat Oncol Biol Phys 2009;74:355-362.

[35]   Dehing-Oberije C, De Ruysscher D, van Baardwijk A, Yu S, Rao B, Lambin P. The importance of patient
       characteristics for the prediction of radiation-induced lung toxicity. Radiother Oncol 2009;91:421-426.

[36]   Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical
       images using advanced feature analysis. Eur J Cancer 2012;48:441-446.

[37]   Vaidya M, Creach KM, Frye J, Dehdashti F, Bradley JD, El Naqa I. Combined PET/CT image character-
       istics for radiotherapy tumor response in lung cancer. Radiother Oncol 2012;102:239-245.

[38]   Thorwarth D, Alber M. Implementation of hypoxia imaging into treatment planning and delivery.
       Radiother Oncol 2010;97:172-175.

Chapter 4

OPEN-SOURCE DATA-SHARING INFRASTRUCTURE

International Data-sharing for Radiotherapy Research: an open-source based infrastructure for multicentric clinical data mining

Erik Roelofs*, André Dekker*,
Elisa Meldolesi, Ruud G.P.M. van Stiphout,
Vincenzo Valentini*, Philippe Lambin*; *Equal contribution

Abstract

Extensive, multifactorial data sharing is a crucial prerequisite for current and future (radiotherapy) research. However, the cost, time and effort to achieve this are often a roadblock. We present an open- source based data-sharing infrastructure between two radiotherapy departments, allowing seamless exchange of de-identified, automatically translated clinical and biomedical treatment data.

Introduction

Sharing data across institutions is required for multi-institutional radiotherapy research [1]. Besides exchanging data for specific research projects, there is a recognized need to establish a culture of data pooling both within the radiotherapy [2] and the broader cancer community [3]. For the transition from population based treatment options (where "one size fits all") towards personalized medicine we are increasingly depending on decision support systems that require large heterogeneous datasets [4–8]. Randomized controlled trials hardly offer such data with only 3% of adult cancer patients included in trials [9–11]. However, aggregating routinely collected real-time biomedical patient data and innovative "rapid-learning" research techniques allow us to use the knowledge of the masses for the benefit of the individual [3,12,13].

Medical informatics driven research, for instance in the field of predictive modelling, requires a large amount of data to provide sufficient statistical power to act as acceptable decision supporting tools. Furthermore, another substantial amount of data is needed for validation of the models, preferably by external datasets.

This brings up some stringent and challenging demands on the quality as well as the quantity of the data. Data of inferior quality do not improve by pooling it with other data. It can actually worsen the value of good datasets. Work by the Quality Assurance Review Center (QARC; www.qarc.org) and a review by the EORTC show the importance of proper Quality Assurance (QA) programmes in collaborative efforts and the long history thereof in the field of Radiotherapy [14–16].

Another challenge for data-sharing initiatives in the field of biomedical research is that the investigated data are often multifactorial, comprising of laboratory data, diagnostic and clinical imaging and treatment outcome data, among others. Combining these data securely can be troublesome, even when sharing between departments within the same institution, let alone when between institutions, especially international ones.

Furthermore, in many research projects, dedicated data management staff are required to translate and copy data into trial-specific case report forms and/or dedicated IT staff are needed to de-identify DICOM images or build databases that are suitable for machine learning and data mining techniques. However, without dedicated staff, the sheer amount of time it takes to collect,

de-identify and share data often is a roadblock to participation in clinical research. With many research projects not or underfunded, especially in the initiation phase, one requires existing staff to balance other duties with these research requests. This causes the process of data sharing to take a long time, despite the cooperation and willingness of everyone involved.

In this technical report, we describe one way to quickly build a low cost, infrastructure that makes sharing of data easier between two institutions wishing to work together, but having different IT systems. This infrastructure was implemented to share data from the Policlinico Universitario Agostino Gemelli in Rome, Italy (Gemelli) to the MAASTRO Clinic in Maastricht, the Netherlands (MAASTRO) to facilitate research projects such as the Thunder clinical trial (NCT00969657, clinicaltrials.gov) and "knowledge engineering" research using data-mining and machine learning techniques to develop predictive models for various cancer sites (www.predictcancer.org).

Material and Methods

Clinical data sources

In general, radiotherapy research requires various types of information:

- Clinical data (e.g. demographics, TNM-stage, date of diagnosis, histopathology, etc.)
- Diagnostic imaging data (e.g. diagnostic and follow-up PET, CT and MR imaging)
- Radiotherapy treatment planning data (e.g. delineation, planning-CT, dose matrix, beam setup, prescribed dose and fractions)
- Radiotherapy treatment delivery data (e.g. cone beam CTs, Orthogonal EPID imaging, delivered fractions)
- Non-radiotherapy treatment data (e.g. surgery, chemotherapy)
- Outcome data (e.g. survival, local control, toxicity).

Typically, in a radiotherapy department, this information is scattered across a number of data sources from a variety of vendors, which do not necessarily share the same patient identification number. In the case of Gemelli, the data sources were as given in Table 1.

Table 1: Data sources at Gemelli

Data type	Data format	Database	Department	Name, Vendor
Clinical Outcome Non-RT treatment	Text	SQL database	Multiple	Spider, Opengraph (local development)
Diagnostic imaging	DICOM-CT DICOM-MR DICOM-PT	DICOM server	Radiology	Careview, Kodak
	DICOM-PT	Optical Disks	Nuclear Medicine	PET workstation, Philips
RT treatment planning	DICOM-CT DICOM-RTDOSE DICOM-RTIMAGE DICOM-RTPLAN DICOM-RTSTRUCT	DICOM server	Radiotherapy	Aria, Varian
RT treatment delivery	Text DICOM-RTRECORD DICOM-RTIMAGE	Sybase Database DICOM server	Radiotherapy	Aria, Varian

Data model

For the research database (DB), a patient-centric data model was designed that would enable queries for both medical data and the existence of imaging data (Figure 1). A simple data model was deliberately chosen to allow easy identification of core disease characteristics but with most information in the form of lists of performed procedures as well as performed imaging study and series.
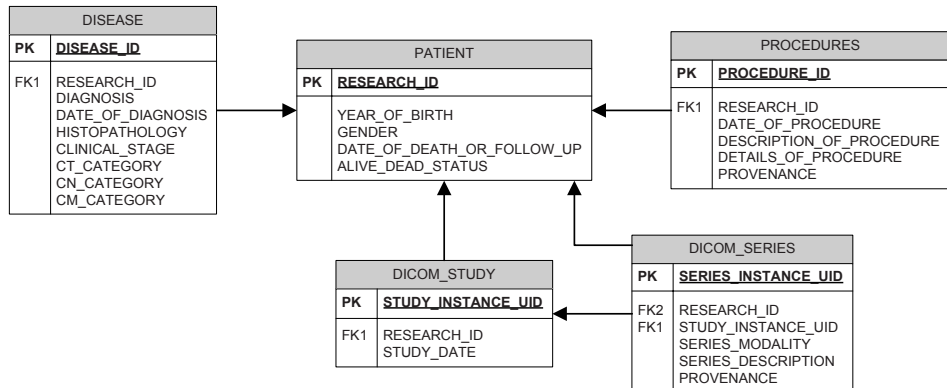
Figure 5: Data model of the research database (ResearchDB). (PK = Primary Key; FK = Foreign Key)

De-identification

A coding scheme was employed in which a secure database is maintained that holds the link (Key) between a unique random patient identification code (ID) and all directly identifying data (ID's as used in the clinical data sources, name, birth date etc.). This Key is maintained by local hospital personnel and only accessible from within the firewall of the hospital. In the research DB, the patient is only identified by the research ID. Rather than applying an irreversible anonymization method to the patient data, we used "coding" or "pseudo-anonymization" to enable extending the datasets with additional information at a later stage, which would otherwise be impossible to do.

Some data elements were not de-identified as they were considered to be important for the research while they only carry a small risk of identifying a patient. The de-identification scheme was reviewed and approved by the local ethical authority. The elements that were kept were: exact dates of various procedures (including treatments), exact dates of diagnosis & death, DICOM UIDs and CT and MR imaging of the head.

Clinical terms and translation

To convert Italian to English standard terms, SNOMED Clinical Terms [17] were used as the dictionary. SNOMED CT is considered as the most comprehensive multilingual medical terminology in the world. A separate database was maintained in which local terms were mapped to the SNOMED-CT concepts and both the preferred term and the concept ID were stored.

Research hardware & software

On a research workstation (Windows 7 64-bit, Intel Xeon, 2.53GHz, 4GB RAM) the following software was installed: SQL Server 2008 (free Express version, Microsoft, Redmond, WA); Clear Canvas Image Server and Workstation (both free and open source, Clear Canvas, Toronto, Canada); DCMTK DICOM toolkit (free and open source, Offis, Oldenburg, Germany); RSNA Clinical Trial Processor (CTP) (free and open source, RSNA, Oakbrook, IL) and Matlab Compiler Runtime (MCR) engine (free, Mathworks, Natick, MA).

Clear Canvas Image Server is a PACS and was installed with a temporary partition holding identifiable DICOM headers and a research partition holding only de-identified DICOM objects. The Clear Canvas Workstation was used for

DICOM import of the Nuclear Medicine department's optical disks. SQL Server 2008 was used to host the mentioned databases as well as the database of the Clear Canvas Image Server.

Finally, for data synchronization a variety of SQL scripts was designed and run through the command line interface. The MCR engine was used to run compiled custom Matlab code in which the DCMTK toolkit was called. The CTP package was used to build a de-identification pipeline (DICOM import → de-identification → DICOM Export) and a file export pipeline (DICOM Import → File Export to shared directory). CTP allows customizable de-identification settings through a web interface or by directly editing an XML configuration file.

Synchronization mechanism – Text Data

The name, birth date, hospital identification number and identification numbers for all data sources were inserted using an SQL query into the secure Key DB for all patients with lung or rectum cancer that had no entry yet. Upon insertion, the patient received a unique research ID, which is a random, non-sequential 8-digit number. Note that the Key DB is a permanent database and that patients will keep the new ID assigned to them.

A second SQL query mapped local terms for demographics, disease characteristics, procedures and clinical findings to a SNOMED term using the dictionary and inserted this information into the Patient, Disease, Procedure and Clinical Finding tables of the research DB.

The above two SQL queries were scheduled to run nightly as to minimize interference with the clinical work, although testing did not indicate any effect of these queries on the clinical process.

Synchronization mechanism – DICOM Data

In compiled Matlab code, a query on the Key DB resulted in a list of patient identification numbers for the DICOM servers. Using DCMTK commands, a so-called DICOM C-FIND request was then sent to each DICOM server, which requested the unique identifiers for all series for a given patient being of the CT, PT, MR, RTPLAN, RTSTRUCT, RTDOSE, RTRECORD or RTIMAGE modality. In Matlab, these DICOM data were compared to the existing data in the research PACS. When the data was not yet available, the Matlab code issued a DICOM C-MOVE request to transfer data from the DICOM Server to the private temporary partition of the research PACS.

In a similar manner, a separate Matlab executable compared DICOM data between the temporary and research PACS. For new datasets, a C-MOVE request was issued to transfer these from the temporary PACS to the CTP de-identification pipeline, de-identified the data and stored them in the research PACS. The data synchronization process is depicted in Figure 2.
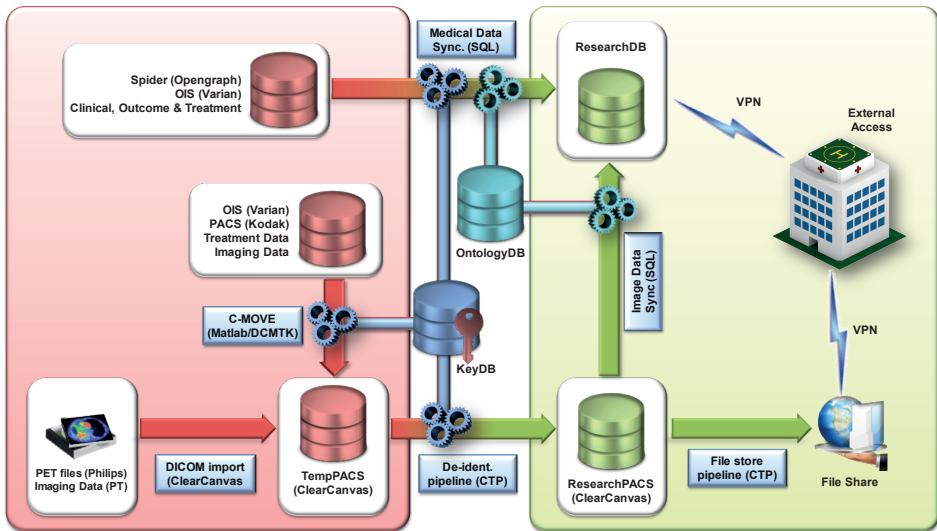
Figure 6: Overview of data sources, flow and external access. Terms as mentioned in the table below:

Term	Meaning
CTP	Clinical Trial Processor
KeyDB	Secure key database linking the random research patient identifier (ResearchID) and original patient data
OntologyDB	Database storing preferred terms and the concept IDs after mapping local terms to the SNOMED CT concepts.
ResearchDB	Research database holding medical data and imaging meta data
ResearchID	Random patient identification code
ResearchPACS	Research PACS partition holding only de-identified DICOM objects
SeriesInstanceUID	Unique series identifier for all images in a series for a given patient
SNOMED-CT	Systematized Nomenclature of Medicine - Clinical Terms
TempPACS	Temporary PACS partition holding identifiable DICOM headers
UID	Unique IDentifier

Data retrieval

To securely transfer data from the research DB and PACS to a remote location, first a VPN connection was established. VPN was chosen, as this is the *de-facto* standard to create a so-called "tunnel" (a secured one-to-one connection via the public Internet). An SQL query was then issued to the research DB (and only this DB). For the DICOM data, direct retrieval was not possible due to security concerns. Instead, a C-MOVE request had to be sent (e.g. using DCMTK commands) to send images from the research PACS to a CTP file-share pipeline. This pipeline temporarily stored the requested images in a shared location from where the images could be copied.

Legal and ethical

A collaboration and data transfer agreement was signed which describes the type of data, the permitted use and the protection of the data. This agreement was submitted to and approved by the local ethical authority. An example agreement can be found online in the supplemental data of this article at www.cancerdata.org/10.1016/ j.radonc.2013.11.001.

Results and Discussion

In this report, we have described one way to build a data-sharing infrastructure quickly and at low cost, which makes sharing of data easier between two institutions wishing to collaborate and exchange data. Starting with contractual arrangements on the use and protection of the data, it is possible to build a data infrastructure that uses free or open source products and that can address common problems such as data de-identification, the need to share DICOM RT objects and the need for a common language/ontology.

The data-sharing infrastructure was implemented in two months (mid 2010). As of July 2012 already, the research DB holds about 1.500 lung, 1.500 rectum, 1.000 head & neck, 300 cervix and 200 pancreas cancer patients and the research PACS holds about 13.500 de-identified DICOM studies using a total of 3 TB storage.

All new rectum and lung cancer patients are added automatically to the research DB and PACS through the nightly synchronization, with a projected yearly increase of 500-600 patients. With a Good Clinical Practice (GCP) compliant [18] data warehouse solution that was previously reported in a unicentric

setting, these databases could easily provide high quality dose-volume data, for instance [19]. This would significantly reduce the amount of time needed by data managers to support existing or new clinical trials, especially in a multicentric setting.

Unexpected problems were that patients were known by different IDs in the various information systems. One reason being that department-specific vendors pose different restrictions (e.g. no leading zero's, or limit on the number of digits) so that the central hospital ID could not be applied. Furthermore, information from external hospitals was often identified by the external hospital patient ID. When personal identification was achievable, for instance when patients brought images on DVD from another hospital, the data were combined, but linking based on name or birth date alone was not done. The level of detail when removing identifiable information or the decision to maintain this is a matter of the local ethics committee and should be considered per study cohort.

The infrastructure has been used to pool two existing datasets from both institutes [20] and is actively used in the Thunder trial (NCT00969657, ClinicalTrials.gov) in which multiple PET/CT scans before, during and after concurrent chemo-radiation are used to predict the likelihood of a complete pathological response in rectum cancer patients [21]. A second NIH funded trial using the infrastructure is the "Radiomics" [22] trial (NCT01302626) in which advanced analysis of CT and PET images of lung cancer patients are correlated with genetic expression and outcome.

Another active application of the data-sharing infrastructure is "*knowledge engineering*" research. In this type of research, large datasets are required in combination with machine learning techniques to learn predictive models for outcome in cancer patients such as we described before [7,23,24]. To prove the accuracy of the model and prevent erroneous conclusions due to over-fitting external validation against independent datasets is crucial. The reported framework facilitated such validation for published predictive models in rectum cancer [5,20].

Finally, the infrastructure is used for the initiation phase of prospective clinical trials to test hypotheses on historical data or to estimate the number of patients that might be included in a new trial given a certain patient selection. Examples showing the power of the infrastructure are a) outcome and number of patients older than 75 treated radically with limited stage small lung cancer, b) patients with a PET/CT scan before and during a course of radiotherapy of lung cancer or c) occurrences of dry mouth in a cohort of 1000 patients being treated with conformal versus intensity-modulated radiotherapy.

Our infrastructure has shown to solve the issue of sharing large amounts of medical and imaging data for research purposes. The automatic transfer from the clinical data flow to a multicentric research environment facilitates daily aggregation of information and offers valuable data for mining and validation. The basic infrastructure is easily extended with more tumour sites or additional relevant procedures. These extensions require modifications of the nightly synchronization and mapping table, which only take minutes. In a manual data sharing method, adding a data element or disease site would require dedicated staff going through all (paper) charts again twice according to GCP guidelines.

The presented method should be seen as an *ad-hoc* data sharing infrastructure and an initial step to collaborate with two institutions. Previously, we described the MISTIR framework for *in silico* clinical trials in Radiotherapy, which enables multiple institutions to securely exchange research data under a strict protocol and QA programme [25]. One example of a QA measure we applied for MISTIR as well as for the Thunder trial is to calibrate the standardized uptake values (SUV) from the PET images to enable reliable SUV-driven segmentation among different institutions [26].

In a joint initiative from the Center for Translational Molecular Medicine (CTMM), the Dutch Cancer Society, the Dutch Heart Foundation, the Netherlands Federation of University Medical Centers, the Netherlands Bioinformatics Centre, the String of Pearls Initiative (PSI) and the Netherlands eScience Center (NLeSC), among others, a medical informatics project called Translational Research IT (TraIT) was started to support the transition of highly promising medical research of twenty-one CTMM projects into clinical care (www.ctmm-trait.nl). Within the TraIT program, we have launched the national BioMedical Imaging Archive (www.bmia.nl) to store and exchange clinical imaging data using the open-source grid-based caBIG™ tools offered by the NCI, NIH [27–29]. Furthermore, TraIT has set up a national instance of OpenClinica (OpenClinica LLC, Waltham, USA) to support electronic data capture for large inter-institutional trials among most of the university medical centres (www.openclinica.nl). Efforts are currently undertaken to offer direct image access from the electronic case report forms.

We hope that prototyping projects such as MISTIR and the one described here and extensions such as TraIT and the advanced federated EuroCAT computer network (www.eurocat.info) with participating institutions in Belgium, Germany and The Netherlands will contribute to global initiatives that offer data-sharing and research capabilities for the entire oncology sector.

Acknowledgement

Appendix: Supplemental Data

To support the open source and open data concept, we published supplementary data online at www.cancerdata.org/10.1016/j.radonc.2013.11.001.

References

[1] Boulton G, Rawlins M, Vallance P, Walport M. Science as a public enterprise: the case for open data. The Lancet 14;377:1633–5.

[2] Deasy JO, Bentzen SM, Jackson A, et al. Improving Normal Tissue Complication Probability Models: The Need to Adopt a "Data-Pooling" Culture. Int J Radiat Oncol Biol Phys 2010;76:S151–S154.

[3] Abernethy AP, Etheredge LM, Ganz PA, et al. Rapid-Learning System for Cancer Care. J Clin Oncol 2010;28:4268 –4274.

[4] Lambin P, Petit SF, Aerts HJWL, et al. The ESTRO Breur Lecture 2009. From population to voxel-based radiotherapy: Exploiting intra-tumour and intra-organ heterogeneity for advanced treatment of non-small cell lung cancer. Radiother Oncol 2010;96:145–52.

[5] Valentini V, van Stiphout RGPM, Lammering G, et al. Nomograms for Predicting Local Recurrence, Distant Metastases, and Overall Survival for Patients With Locally Advanced Rectal Cancer on the Basis of European Randomized Clinical Trials. J Clin Oncol 2011;29:3163 –3172.

[6] Egelmeer AGTM, Velazquez ER, de Jong JMA, et al. Development and validation of a nomogram for prediction of survival and local control in laryngeal carcinoma patients treated with radiotherapy alone: A cohort study based on 994 patients. Radiother Oncol 2011;100:108–15.

[7] Dehing-Oberije C, De Ruysscher D, Petit S, et al. Development, external validation and clinical useful-ness of a practical prediction model for radiation-induced dysphagia in lung cancer patients. Radiother Oncol 2010;97:455–61.

[8] Valentini V, Lambin P, Myerson RJ. Is it time for tailored treatment of rectal cancer? From prescribing by consensus to prescribing by numbers. Radiother Oncol 2012;102:1–3.

[9] Movsas B, Moughan J, Owen J, et al. Who enrolls onto clinical oncology trials? A radiation Patterns Of Care Study analysis. Int J Radiat Oncol Biol Phys 2007;68:1145–50.

[10] Murthy VH, Krumholz HM, Gross CP. Participation in cancer clinical trials: race-, sex-, and age-based disparities. JAMA 2004;291:2720–6.

[11] Grand MM, O'Brien PC. Obstacles to participation in randomised cancer clinical trials: a systematic review of the literature. J Med Imaging Radiat Oncol 2012;56:31–9.

[12] Lambin P, van Stiphout RGPM, Starmans MHW, et al. Predicting outcomes in radiation oncology— multifactorial decision support systems. Nat Rev Clin Oncol 2012;10:27–40.

[13] Lambin P, Roelofs E, Reymen B, et al. "Rapid Learning health care in oncology" – An approach towards decision support systems enabling customised radiotherapy. Radiother Oncol 2013;109:159–64.

[14] Peters LJ, O'Sullivan B, Giralt J, et al. Critical Impact of Radiotherapy Protocol Compliance and Quality in the Treatment of Advanced Head and Neck Cancer: Results From TROG 02.02. J Clin Oncol 2010;28:2996–3001.

[15] Weber DC, Poortmans PMP, Hurkmans CW, et al. Quality assurance for prospective EORTC radiation oncology trials: The challenges of advanced technology in a multicenter international setting. Radiother Oncol 2011;100:150–6.

[16] Weber DC, Tomsej M, Melidis C, Hurkmans CW. QA makes a clinical trial stronger: Evidence-based medicine in radiation therapy. Radiother Oncol 2012;105:4–8.

[17] International Health Terminology Standards Development Organization. SNOMED Clinical Terms (SNOMED CT) n.d.

[18] European Medicines Agency - Inspections - Good-clinical-practice compliance 2013.

[19] Roelofs E, Persoon L, Nijsten S, et al. Benefits of a clinical data warehouse with data mining tools to collect data for a radiotherapy trial. Radiother Oncol 2013;108:174–9.

[20] Van Stiphout RGPM, Lammering G, Buijsen J, et al. Development and external validation of a predictive model for pathological complete response of rectal cancer patients including sequential PET-CT imag-ing. Radiother Oncol 2011;98:126–33.

[21] Janssen MHM, Öllers MC, Riedl RG, et al. Accurate Prediction of Pathological Rectal Tumor Response after Two Weeks of Preoperative Radiochemotherapy Using 18F-Fluorodeoxyglucose-Positron Emission Tomography-Computed Tomography Imaging. Int J Radiat Oncol Biol Phys 2010;77:392–9.

[22] Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441–6.

[23] Dehing-Oberije C, Yu S, De Ruysscher D, et al. Development and External Validation of Prognostic Model for 2-Year Survival of Non–Small-Cell Lung Cancer Patients Treated With Chemoradiotherapy. Int J Radiat Oncol Biol Phys 2009;74:355–62.

[24] Dehing-Oberije C, Aerts H, Yu S, et al. Development and Validation of a Prognostic Model Using Blood Biomarker Information for Prediction of Survival of Non–Small-Cell Lung Cancer Patients Treated With Combined Chemotherapy and Radiation or Radiotherapy Alone (NCT00181519, NCT00573040, and NCT00572325). Int J Radiat Oncol Biol Phys 2011;81:360–8.

[25] Roelofs E, Persoon L, Qamhiyeh S, et al. Design of and technical challenges involved in a framework for multicentric radiotherapy treatment planning studies. Radiother Oncol 2010;97:567–71.

[26] Öllers M, Bosmans G, van Baardwijk A, et al. The integration of PET-CT scans from different hospitals into radiotherapy treatment planning. Radiother Oncol 2008;87:142–6.

[27] Langella S, Oster S, Hastings S, et al. The Cancer Biomedical Informatics Grid (caBIG^{TM}) Security Infrastructure. AMIA Annu Symp Proc 2007;2007:433.

[28] Oster S, Langella S, Hastings S, et al. caGrid 1.0: An Enterprise Grid Infrastructure for Biomedical Research. J Am Med Inform Assoc 2007;15:138–49.

[29] Buetow KH. An infrastructure for interconnecting research institutions. Drug Discover Today 2009;14:605–10.

Chapter 5

MISTIR: Multicentric Collaboration Framework

*Design of and technical challenges involved in a framework
for multicentric radiotherapy treatment planning studies*

Erik Roelofs, Lucas Persoon, Sima Qamhiyeh, Frank Verhaegen,
Dirk De Ruysscher, Michael Scholz, Gheorghe Iancu, Martijn Engelsman,
Coen Rasch, Lambert Zijp, Gert De Meerleer, Marc Coghe, Johannes Langen-
dijk, Cornelis Schilstra, Madelon Pijls-Johannesma, Philippe Lambin

Abstract

This report introduces a framework for comparing radiotherapy treatment planning in multicentric *in silico* clinical trials. Quality assurance, data incompatibility, transfer and storage issues, and uniform analysis of results are discussed. The solutions that are given provide a useful guide for the set-up of future multicentric planning studies or public repositories of high quality data.

Introduction

In radiation oncology, *in silico* clinical trials or planning studies are increasingly popular for investigating different treatment options without harming the patient [1-3]. They can be described as modelling studies that, based on virtual patient material (imaging), validated procedures (including *a priori* sample size calculation) and models, provide an alternative method of exploring or generating hypotheses. The modelling results should then be validated in real life with prospective clinical trials[4].

The concept of treatment planning comparison studies is not new. Starting in the early 80s, the National Cancer Institute funded various projects to evaluate and compare different treatment modalities [5,6]. Back then, a lot of tools and procedures we now take for granted (e.g. 3D dose evaluation, multi-modal image registration, plan optimisation) were missing or insufficiently implemented. This hindered the proceedings of those projects. In the three decades that followed a lot of these issues were overcome. Although the comparison of radiotherapy treatment options in a multicentric setting is still challenging it is much less problematic as in the early days.

One such multicentric *in silico* clinical trial that is based on the MISTIR (acronym for "Multicentric *In Silico* Trials In Radiotherapy") framework presented here is ROCOCO (**R**adiation **O**ncology **C**ollaborative **C**omparison) [7]. This is an emulation of clinical trials in photon, proton and heavier particle radiotherapy for tumours with high incidence. It comprises the treatment planning (TP) comparison of lung, prostate and head and neck tumours for 25 patients in each group. Currently, eight international institutes are performing TP for conventional and intensity-modulated photon, passive scattered, scanning or intensity-modulated proton and carbon ion radiotherapy. MAASTRO CLINIC serves as the coordinating Data Centre (DC) as well as being a participant.

During the course of the project various organisational and technical issues arose and were solved, thus maturing the framework into its current state (Fig. 1). As we believe the set-up is an *all-in-one* solution that can be used for a wide range of *in silico* clinical trials in radiotherapy, planning studies and/or for public repository of anonymised datasets, our intention with this technical note is to provide a step-by-step guide for building such a framework. All major action blocks will be discussed and critical issues on data-exchange will be explored. Information that is kept up-to-date can be found online on the MISTIR website: www.mistir.info.
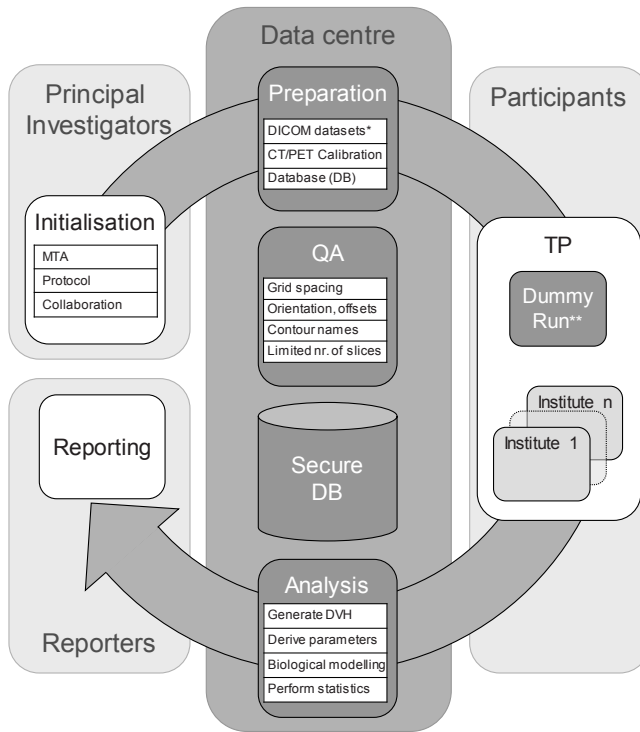
Figure 1 A diagram illustrating the major actions of a project that is based on the currently presented in silico trial framework. After uploading data to the DB it is checked for integrity and validity by various QA procedures (some items shown). The blocks surrounded by the "Data centre" frame are performed centrally to avoid data corruption and minimise uncertainty due to the use of different algorithms.

*DICOM images can originate from multiple institutions. However, before allowing other participants to use them the DC performs the initial QA procedures.

**The Dummy Run is a joint action by the DC and participants as part of the QA programme.

Methods

For an international, multicentric *in silico* clinical trial that started in January 2008 a framework was built using five major action blocks (Figure 1). A centrally hosted database (DB) was designed to host the protocol, the patient datasets, the TP results and the analysis thereof. Secure access to the database was granted to authorised project participants by means of the FTPS (secured File Transfer Protocol) and HTTPS (secured Hyper Text Transfer Protocol) protocols

[8,9]. We used Public Key Infrastructure (PKI) certificates (the gold standard security method in data exchange of medical data [10-12]) to verify the user's identity before login was allowed.

To facilitate data sharing, the participants agreed to use the iternationally accepted Digital Imaging and Communications in Medicine standards (DICOM) [13]. In the event of a participant being unable to comply, the RTOG (Radiation Therapy Oncology Group) format [14] was accepted.

The five action blocks of the framework, as described below, were: initialisation, preparation, treatment planning, analysis and reporting.

Initialisation

Following the template of a real clinical trial, a protocol was set up in which the trial hypothesis, the sample size and the methodology was described. The structure and management of the project was defined by project goals and milestones. With respect to the topic that was studied (the comparison of treatment techniques or modalities) the protocol described the study endpoints, differentiating primary and secondary endpoints and how differences were analysed. Furthermore, it described the available data and means to access it in a secure manner. A Material Transfer Agreement (MTA) was formulated in which the participants' roles, rights and liabilities were described. After signing the MTA, the participant was granted access to the central database. Templates of the protocol and MTA are available online.

Preparation

With the prepared protocol, eligible patients were accrued and their datasets were gathered. By consecutively including patients that satisfied the protocol, no selection bias was introduced. A set of CT images and delineations of targets and critical organs in the DICOM RT STRUCT format were de-identified and uploaded to the database. Furthermore, fluor-18-fluorodeoxyglucose (FDG) PET and MR images were added. For the available 4D CT/PET data, the 3D motion vector of the target was determined and included in the dataset [2]. Because the FDG-PET images were used to automatically delineate high-uptake regions with a method validated with pathology [15-17], the standardised uptake value (SUV) conversion curves of data-supplying institutes were calibrated [18]. This guaranteed that the auto-delineation algorithms used in the data-supplying institutes created equivalent contours.

Treatment Planning

To perform TP, the centres downloaded the datasets from the database. It was mandatory for each centre to test if any shifts occurred in the structure sets during data import. To validate the import, screen captures in JPG format of representative slices were uploaded for verification. The centre then proceeded with TP according to the strict criteria as described in the project protocol. The TP results were exported by the participants in DICOM RT format and uploaded to the DB.

Analysis

In order to analyse the results and compare TP modalities, the DC used the CT, the delineations and the dose matrices stored in the DB. Visualisation of the TP results was done using MATLAB (The MathWorks, Natick, Massachusetts), CERR [19] or VODCA [20].

For the calculation of dose-volume histograms (DVHs), user-developed MATLAB code was used instead of requesting DVH parameters from the participants. The centrally performed calculations reduce any uncertainties that can arise due to differences in the algorithms of the various treatment planning systems (TPSs).

The DVHs were then used to derive relevant dose metrics of the different structures such as mean, maximum and minimum dose, conformity indices (CI), inhomogeneity coefficient (IC) and tumour control probability and normal tissue complication probability (TCP and NTCP respectively) where models were available. The abovementioned metrics were calculated using the strict formulation described in the project protocol. The results were made available by entering them into the DB.

Reporting

Finally, the validated data in the DB was available for publication by the designated project members.

Quality Assurance

A large collaborative project such as ROCOCO needs a well-designed data exchange set-up. Data supplied to all participants originates from multiple sources including CT, PET and MR scanners and from the TPS. TP will be based on the data available in the DB and participants will upload their data from their systems to the DB. Currently, TP has been performed using XiO/Focal (CMS Software, Elekta), Pinnacle (Philips), Virtuos/TRiP (in-house: DKFZ Heidelberg / GSI Darmstadt) and HIPLAN (in-house: NIRS). With such a variety of programs involved and due to the fact that the DICOM standard allows vendors to incorporate proprietary information, interoperability problems easily arise with data import/export [21]. Therefore, it is essential to define a proper Quality Assurance (QA) programme.

At MAASTRO CLINIC, an in-house developed program called *DIGITrans* is used in daily clinical routine that normalises all DICOM (RT) data to prevent interoperability issues and guarantees data consistency. This includes matching the patient demographics of exported data from the scanners, TPS or record-and-verifying system to the electronic health record. Furthermore, we "brand" the images by invisibly adding referencing DICOM tags to the image data.

During the set-up of the MISTIR framework many issues were addressed (Table 1) for which we reused part of the *DIGITrans* procedures or developed new ones. The MISTIR QA procedures now prevent and correct problems with:

- SUV (PET) calibration
- CT calibration (Hounsfield unit to electron density conversion)
- correct media assignment, if needed (e.g. in Monte Carlo dose calculations)
- DICOM (RT) data connectivity
- de-identification
- 4D data phase consistency
- orientation / transformation errors
- number of slices
- contour names
- contour interpolation
- multiple contours on a single slice
- invalid dose grid sizes
- other protocol violations

Table 2 A selection of problems, solutions and remarks on the implementation of the MISTIR QA procedures[3].

Topic	Problems occurred during dummy-run	Solution	Remarks
De-identifi-cation	The TPS, discards breathing phases in 4D datasets, due to reduced DICOM references in TPS export.	Write code to directly modify the needed tags.	We use *dcmodify* from DCMTK[4].
	The TPS does not handle PET or MR export.		
Integrity	TPS changes coordinate systems during DICOM import/export.	Check for transformations whenever data is entering another system.	We use screenshots to visually check for translations, rotations, inversions etc.
	TPS drops the *Frame of Reference UID* during DICOM import/export.	Encode the *Frame of Reference UID* into the CT image data.	We use the offset vector to restore transformed datasets.
Slice thickness	TPS cannot handle CT images and contours with varying slice thickness.	Resample the CT data, re-import into TPS and use the interpolation function to generated contours on new slices. Manually correct contours (e.g. in case of bifurcations)	
	TPS cannot handle difference in reconstruction thickness and slice spacing.	Make the slice thickness equal to the difference between the slice locations.	Radiologically this is not the same, however, it can be ignored for radiotherapy.
Nr. of slices	TPS could not handle >99 CT slices.	Make a subset of the data or resample if critical areas would be cut off.	Add ~5cm margin around beam edges to account for patient scatter.
Structure names	TPS could not handle structure names starting with a number.	Change the structure name to begin with a letter.	
Internet protocols	Secured network transfer between institutes using FTPS not possible as the protocol is often blocked by hospital firewalls.	Use an HTTPS as well as an FTPS server.	We use DRUPAL[5] and FileZilla Server[6], respectively. We use PKI certificates to guarantee security.
File management	An HTTPS server does not always allow easy file manipulation.	Add a module for file management.	We use the WebFM module[7].
Data corruption	With many files in large datasets, upload is prone to errors.	Pack the datasets before upload.	We propose checksums to verify data (e.g. MD5 or SHA1).

---

[3] Refer to MISTIR: http://www.mistir.info for more details

[4] OFFIS Dicom Toolkit: http://dicom.offis.de

[5] DRUPAL: http://drupal.org

[6] FileZilla: http://filezilla-project.org

[7] WebFM: http://drupal.org/project/webfm

To test the set-up of the protocol, DB and QA procedures, first a dummy run is performed with a representative sample of the dataset. The participants are asked to perform TP according to the protocol and send their results to the DB. Upon receiving the datasets, the QA procedures are applied to ensure consistency and validity of the data. In cases where the TP data does not conform to the protocol, the participant is asked to adapt its procedures and redo TP. Then the data is checked again for validity. After acceptance, the participant is asked to proceed with the entire dataset.

Results

The presented MISTIR framework has been successfully used for a multicentric, *in silico* clinical trial that is currently being conducted by the ROCOCO consortium. Twelve institutes are participating and several more have expressed interest in the study.

To ensure data integrity, a series of QA tests is performed manually on any item (including that from the host institute) that is uploaded to the DB before releasing it to the project members. The QA procedures are designed to detect and correct issues regarding the readability of data, compatibility with DICOM (RT) standards and transformations in structure delineations and/or in image coordinates.

The design of the framework has successfully been tested for several different sectors of the trial. Implementation of the mentioned procedures has successfully led to TP results for prostate and lung cancer [22,23].

Discussion

We have demonstrated that building a functional data management and analysis framework for an international, multicentric *in silico* clinical trial is feasible. Its QA procedures are able to detect data inconsistencies and prevent incorrect data analysis.

In its current state, MISTIR is used for file-based data exchange and manual data analysis. To support large multicentric trials with improved data transfer and warehousing of the DICOM RT objects and data mining capabilities we are working on implementing a DICOM RT compliant DB (PACS). To ensure semantic and technical interoperability, recommendations from standardising

initiatives for healthcare and clinical research such as caBIG [24], CDISC [25] and IHE [26] are taken into account.

One application that would strongly benefit from the availability of large validated, multicentric datasets is the modelling of radiotherapy outcome [27,28]. For satisfying significance of the predictive models a large number of datasets is necessary, while for robustness it is best to validate the models on foreign datasets. We believe that public sharing of these datasets would stimulate this research on a global scale, as it already has done in genomics research where public data repositories are common practice [29]. Whether a centralised set-up based on the current MISTIR platform is needed or that it should evolve into a de-centralised, grid-based solution such as the GridCAD and GridIMAGE initiatives for radiology [30,31] remains to be determined. This is considered to be beyond the scope of this paper. A very recent series of papers of the QUANTEC consortium summarised the currently available dose/volume/outcome data for normal tissues [32]. In a vision paper of the same series they concluded that storage of high quality datasets in repositories should become a common strategy [33]. We believe the described technology will help to realise this goal.

Conclusion

We presented MISTIR: a complete and secure framework for *in silico* clinical trials using the output of treatment planning on prepared datasets. It is successfully being used in the *in silico* clinical treatment planning trial that the ROCOCO consortium currently is conducting. By defining strict planning protocols and using validated algorithms, the results of the *in silico* trials can be used as a starting point for validating and/or generating hypotheses, without harming any patient and at a reasonable cost. These hypotheses need further testing in prospective phase II/III trials.

The step-by-step description of the different building blocks and procedures that are used in MISTIR offer a useful guide for the set-up of future *in silico* trials and/or public repositories.

Acknowledgments

References

[1] van Baardwijk A, Bosmans G, Bentzen SM, et al. Radiation dose prescription for non-small-cell lung cancer according to normal tissue dose constraints: an in silico clinical trial. Int J Radiat Oncol Biol Phys 2008;71:1103-1110.

[2] Bosmans G, Buijsen J, Dekker A, et al. An "in silico" clinical trial comparing free breathing, slow and respiration correlated computed tomography in lung cancer patients. Radiother Oncol 2006;81:73-80.

[3] Jones B, Dale RG. Further radiobiologic modeling of palliative radiotherapy: use of virtual trials. Int J Radiat Oncol Biol Phys 2007;69:221-229.

[4] Brada M, Pijls-Johannesma M, De Ruysscher D. Current clinical evidence for proton therapy. Cancer J 2009;15:319-324.

[5] Goitein M, Lyman J, Maor M. "Report of the Working Groups on the Evaluation of Treatment Planning for Particle Beam Radiotherapy". Bethesda, MD: Radiotherapy Development Branch, Research Program Division of Cancer Treatment, National Cancer Institute, 1987.

[6] Evaluation of high energy photon external beam treatment planning: project summary. Photon Treatment Planning Collaborative Working Group. Int J Radiat Oncol Biol Phys 1991;21:3-8.

[7] Qamhiyeh S, Pijls-Johannesma M, Verhaegen F, et al. ROCOCO: An International, Multicentric In Silico Clinical Trial In Lung, Prostate And Head&Neck Cancer: Evaluation Of Feasibility. Third International Conference on Translational Research and Pre-Clinical Strategies in Radiation Oncology 2009;Geneva, Switzerland.

[8] Securing FTP with TLS. The Internet Engineering Task Force (IETF), 2005. (Accessed January 21, 2010, at www.ietf.org/rfc/rfc4217.txt)

[9] HTTP Over TLS. The Internet Engineering Task Force (IETF), 2000. (Accessed January 21, 2010, at www.ietf.org/rfc/rfc2818.txt)

[10] Schutze B, Kammerer M, Klos G, Mildenberger P. The Public-Key-Infrastructure of the Radiological Society of Germany. Eur J Radiol 2006;57:323-328.

[11] Pharow P, Blobel B. Public key infrastructures for health. Stud Health Technol Inform 2003;96:111-117.

[12] Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile. The Internet Engineering Task Force (IETF), 2008. (Accessed January 21, 2010, at www.ietf.org/rfc/rfc5280.txt)

[13] American College of Radiology NEMA. ACR-NEMA digital imaging and communications standard: ACR-NEMA standards publication no. 300-1988. ACR-NEMA, 1988.

[14] Baxter BS, Hitchner LE, Jr GQM. AAPM Report No. 10: A Standard Format for Digital Image Exchange. American Association of Physicists in Medicine, 1982.

[15] Stroom J, Blaauwgeers H, van Baardwijk A, et al. Feasibility of pathology-correlated lung imaging for accurate target definition of lung tumors. Int J Radiat Oncol Biol Phys 2007;69:267-275.

[16] van Baardwijk A, Bosmans G, Boersma L, et al. PET-CT-based auto-contouring in non-small-cell lung cancer correlates with pathology and reduces interobserver variability in the delineation of the primary tumor and involved nodal volumes. Int J Radiat Oncol Biol Phys 2007;68:771-778.

[17] van Baardwijk A, Bosmans G, van Suylen RJ, et al. Correlation of intra-tumour heterogeneity on 18F-FDG PET with pathologic features in non-small cell lung cancer: a feasibility study. Radiother Oncol 2008;87:55-58.

[18] Ollers M, Bosmans G, van Baardwijk A, et al. The integration of PET-CT scans from different hospitals into radiotherapy treatment planning. Radiother Oncol 2008;87:142-146.

[19] Deasy JO, Blanco AI, Clark VH. CERR: a computational environment for radiotherapy research. Med Phys 2003;30:979-985.

[20] VODCA: View-Oriented, Distributed, Cluster-based Approach. (Accessed Februari 24, 2010, at www.vodca.ch)

[21] Henderson M, Behlen FM, Parisot C, Siegel EL, Channin DS. Integrating the healthcare enterprise: a primer. Part 4. The role of existing standards in IHE. Radiographics 2001;21:1597-1603.

[22] Roelofs E, Engelsman M, Rasch C, et al. Results Of A Multicenter In Silico Trial Comparing Photons And Protons For Radiotherapy Of Non-Small Cell Lung Cancer. Radiother Oncol 2009;92;Suppl. 1:S108.

[23] Meerleer GD, Roelofs E, Engelsman M, et al. To compare treatment planning results for primary radiotherapy of prostate cancer using proton, 3-dimensional conformal and intensity modulated radiotherapy. 48th Meeting of the Particle Therapy Co-Operative Group 2009. doi:10.3205/09ptcog051

[24] Fenstermacher D, Street C, McSherry T, Nayak V, Overby C, Feldman M. The Cancer Biomedical Informatics Grid (caBIG™). Conf Proc IEEE Eng Med Biol Soc 2005;1:743-746.

[25] Souza T, Kush R, Evans JP. Global clinical data interchange standards are here! Drug Discov Today 2007;12:174-181.

[26] Siegel EL, Channin DS. Integrating the Healthcare Enterprise: a primer. Part 1. Introduction. Radiographics 2001;21:1339-1341.

[27] Dehing-Oberije C, Yu S, De Ruysscher D, et al. Development and external validation of prognostic model for 2-year survival of non-small-cell lung cancer patients treated with chemoradiotherapy. Int J Radiat Oncol Biol Phys 2009;74:355-362.

[28] El Naqa I, Bradley JD, Lindsay PE, Hope AJ, Deasy JO. Predicting radiotherapy outcomes using statistical learning techniques. Phys Med Biol 2009;54:S9-S30.

[29] Entrez Genome. (Accessed Februari 24, 2010, at www.ncbi.nlm.nih.gov/sites/entrez?db=Genome)

[30] Gurcan MN, Pan T, Sharma A, et al. GridIMAGE: a novel use of grid computing to support interactive human and computer-assisted detection decision support. J Digit Imaging 2007;20:160-171.

[31] Pan TC, Gurcan MN, Langella SA, et al. Informatics in radiology: GridCAD: grid-based computer-aided detection system. Radiographics 2007;27:889-897.

[32] Marks LB, Yorke ED, Jackson A, et al. Use of Normal Tissue Complication Probability Models in the Clinic. International Journal of Radiation Oncology*Biology*Physics 2010;76:S10-S19.

[33] Deasy JO, Bentzen SM, Jackson A, et al. Improving Normal Tissue Complication Probability Models: The Need to Adopt a "Data-Pooling" Culture. International Journal of Radiation Oncology*Biology*Physics 2010;76:S151-S154.

Chapter 6

ROCOCO: *IN SILICO* TRIALS

Results of a multicentric in silico clinical trial (ROCOCO): comparing radiotherapy with photons and protons for non-small cell lung cancer

Erik Roelofs, Martijn Engelsman, Coen Rasch,
Lucas Persoon, Sima Qamhiyeh, Dirk de Ruysscher,
Frank Verhaegen, Madelon Pijls-Johannesma,
and Philippe Lambin; on behalf of the ROCOCO Consortium [8]

Abstract

Introduction: This multicentric *in silico* trial compares photon and proton radiotherapy for non-small cell lung cancer patients. The hypothesis is that proton radiotherapy decreases the dose to and the volume of irradiated normal tissues even when escalating to the maximum tolerable dose (MTD) of one or more of the organs at risk (OAR).

Methods: Twenty-five patients, stage IA-IIIB, were prospectively included. On 4D FDG-PET/CT scans, the gross tumour, clinical and planning target volumes and OAR were delineated. Three-dimensional conformal (3DCRT) and intensity-modulated (IMRT) photon as well as passive scattered conformal proton therapy (PSPT) plans were created to give 70 Gy to the tumour in 35 fractions. Dose (de-)escalation was performed by rescaling to the MTD.

Results: Protons resulted in the lowest dose to the OAR, while keeping the dose to the target at 70 Gy. The integral dose was higher for 3DCRT (59%) and IMRT (43%) than for PSPT. The mean lung dose (MLD) reduced from 18.9 Gy for 3DCRT and 16.4 Gy for IMRT to 13.5 Gy for PSPT.

For 10 patients, escalation to 87 Gy was possible for all three modalities. The MLD and integral dose were respectively 40% and 65% higher for photons than for protons.

Conclusions: The treatment planning results of the ROCOCO trial show a reduction of integral dose and the dose to the OAR when treating with protons instead of photons, even with dose-escalation. This shows that PSPT is able to give a high tumour dose, while keeping the OAR dose lower than with the photon modalities.

Introduction

Despite the advances in lung cancer treatment, it remains the number one cause of death among cancer patients in Europe and North America. Of these, 80% are categorized as non-small cell lung cancers (NSCLC) [1,2]. Different approaches in combining surgery, chemotherapy and radiotherapy to increase tumour control and to lower complication rates are the subject of intense investigation [3].

All over the world, proton and also carbon ion (C-ion) radiotherapy is gaining interest and popularity. In the media, it is sometimes bluntly stated that proton therapy yields the best possible treatment for every cancer patient, the arguments being solely based on the reduction of the dose administered to the healthy tissue due to the physical characteristics of charged particles. Current published comparative planning studies demonstrated the advantage in ballistic properties of particle therapy over conventional photon radiotherapy. However, these results were not sufficiently convincing, since they were often monocentric studies, performed with very limited patient numbers and the treatment plans were mostly not performed according to the current clinical guidelines.

Despite a long history of proton radiotherapy, recent reviews have not shown a clear clinical evidence to implement charged-particle therapy on a large scale, mainly due to the lack of randomised controlled trials (RCTs) [4-7]. However, it was concluded from a comprehensive analysis of the current data that particle therapy (PT) was a promising treatment modality for NSCLC [8-10]. Since the cost of PT is considerably higher than conventional radiotherapy with photons, questions arise about the (cost-) effectiveness of this new technology and the need to whether or not perform RCTs [11,12].

In response to the debate whether RCTs are needed for charged particle treatment [13-19] a multicentric *in silico* clinical trial named ROCOCO (**R**adiation **O**ncology **Co**llaborative **C**omparison) was initiated in 2007. It emulates a real clinical trial comparing photon, proton and C-ion therapy for NSCLC, prostate and head & neck cancer using the collaborative MISTIR (Multicentric *In Silico* Trials In Radiotherapy) framework (www.mistir.info). MISTIR uses a secured central database with virtual patient material and trial protocols comparable to the ones used in real RCTs. A set of quality assurance procedures is available to prevent, capture and solve interoperability issues that may arise during data exchange and analysis [20].

With the use of MISTIR and participation of some of the most experienced institutes in the field, we are able to safely explore possible clinically acceptable

treatment options which may pave the way for prospective RCTs with real patients. To the best of our knowledge, no other study has been performed where fourteen international radiotherapy centres jointly perform such an *in silico* trial[9].

This report summarizes the current results of the dosimetric comparison of three-dimensional conformal photon (3DCRT), intensity-modulated photon (IMRT) and passive scattered proton (PSPT) radiotherapy of the NSCLC lung cohort of the ROCOCO trial. The hypothesis is that particle therapy decreases the dose to and the volume of irradiated normal tissue while maintaining an iso-effective dose to the tumour as well as when escalating dose to the maximum tolerable dose. Consequently, the risk of side effects in the surrounding normal tissue is expected to decrease.

Materials and Methods

Study design

An *in silico* clinical trial was performed and data from MAASTRO containing 4D F[18]-labelled fluorodeoxyglucose (FDG) PET/CT images were de-identified and centrally stored on the secured MISTIR database as reported earlier [20]. All volumes were delineated at MAASTRO and planning criteria were described in a pre-defined protocol. The participating institutes downloaded the datasets and performed treatment planning according to the provided protocol using their own clinical treatment planning system (TPS). By using the clinically commissioned TPSs, the correctness of the dose calculations was assured.

Next, the dose matrices in DICOM RT Dose format were uploaded back to the database. Then, after being checked for consistency, the dose matrices were used to calculate the dose-volume histograms (DVH) to derive the dose metrics for the final analysis. All proton beam dose distributions were corrected to Gy(RBE) by applying a relative biological effectiveness (RBE) factor of 1.1. The Gray unit (Gy) was used for absorbed dose reporting of both modalities.

Patient inclusion

Between October 2007 and June 2008, 25 stage I-III NSCLC patients were included consecutively. No selection was made regarding histology, stage or loca-

---

[9] Refer to the list of ROCOCO consortium members in footnote 8.

tion of the tumour, resulting in a heterogeneous group of patients. An overview of the patient characteristics is given in Table 1. All patients were assumed to have a reasonable lung function (a forced expiratory volume in 1 second [FEV$_1$] ≥ 50% of predicted value and a diffusing capacity of lungs for carbon monoxide [DL$_{CO}$] not corrected for alveolar volume ≥ 50%), so that a maximum mean lung dose (MLD) of 19 Gy could be applied [21].

For all patients, the 4D FDG PET/CT was used to determine the clinical staging of the tumours and individual treatment margins.

Table 1 Patient characteristics of the datasets showing histology, TNM classification, staging, top-top amplitudes of the tumour measured on the 4D PET/CT scan and the derived anisotropic margins for photons.

Patient	Histology	TNM	Stage	Amplitude (mm)			CTV to PTV margin (mm)		
				CC	LR	AP	CC	LR	AP
1	squamous	T3N2M0	IIIA	20	5	10	10	6	8
2	adeno	T2N3M0	IIIB	4	3	9	8	6	8
3	adeno	T2N2M0	IIA	8	3	5	8	6	7
4	squamous	T4N0M0	IIIB	3	2	2	7	6	7
5	squamous	T2N1M0	IIB	8	4	6	8	6	7
6	large cell	T4N2M0	IIIB	5	2	3	8	6	7
7	squamous	T4N0M0	IIIB	14	3	4	9	6	7
8	large cell	T1N2M0	IIIA	17	10	20	10	7	10
9	large cell	T4N3M0	IIIB	17	3	5	10	6	7
10	large cell	T2N2M0	IIIA	8	2	2	8	6	7
11	large cell	T2N3M0	IIIB	13	2	1.5	9	6	7
12	large cell	T1N3M0	IIIB	4.5	3	3	8	6	7
13	large cell	T1N2M0	IIIA	7	2	2	8	6	7
14	adeno	T2N2M0	IB	3	2	2	7	6	7
15	squamous	T2N0M0	IIIA	11	3	4	9	6	7
16	adeno	T4N2M0	IIIB	6	3.5	3	8	6	7
17	large cell	T2N3M0	IIIB	6	2	2	8	6	7
18	squamous	T1N0M0	IA	4	5	6	8	6	7
19	squamous	T2M0N0	IB	1	1	1	7	6	7
20	large cell	T2N0M0	IB	5	2	2	8	6	7
21	large cell	T1N2M0	IIIA	10	7	3	8	6	7
22	NSCLC NOS	T4N3M0	IIIB	4.5	2.5	4	8	6	7
23	adeno	T1N1M0	IIA	5	4	2	8	6	7
24	NSCLC NOS	T4N3M0	IIIB	4.5	6	4	8	6	7
25	large cell	T2N2M0	IIIA	3	1	1.5	7	6	7

Abbreviations: CC = cranial - caudal; LR = left - right; AP = anterior – posterior
Squamous: squamous cell carcinoma; adeno: adenocarcinoma; large cell: large cell carcinoma; NSCLC NOS: non-small cell lung cancer not otherwise specified

Target and OAR definitions

Target volumes, critical organs-at-risk (OAR) and other normal tissues were delineated on each slice conform the literature: tumour [22], lymph nodes [23], OAR (spinal cord, oesophagus, lungs, heart) [24]. The spinal cord was considered to be at the inner margin of the bony spinal canal and was drawn throughout the whole CT scan. For the oesophagus, the contour of the outside muscle wall was followed from the distal end of the larynx to the gastro-oesophageal junction. Contouring of the lungs was done using automatic delineation by thresholding on the treatment planning system with a manual correction where needed. The heart was contoured from the apex to the origin of the large vessels, including the pericardium.

Individual peak-to-peak tumour motion was determined using the 4D CT and delineations of the tumour and OAR were projected on the mid-ventilation phase (50% exhale) [22]. For all patients, a gross tumour volume (GTV) was defined based on the 4D FDG PET/CT data. The GTV was defined as the primary tumour on CT and lymph nodes positive on a PET scan or proven to be positive on mediastinoscopy, transoesophageal or transbronchial puncture. No elective nodal irradiation was performed. Syngo TrueD (Siemens Medical Solutions, Malvern, PA) was used to automatically define the PET positive areas by using a pathology validated source-to-background method [25-27].

The clinical target volume (CTV) was defined as the GTV with a margin of 5 mm. Next, the individual, anisotropic planning target volume (PTV) margins for photons were calculated in three directions (CC, LR and AP) using a margin recipe (refer to the Appendix, Eq. A1). For protons, the margins from CTV to PTV are determined differently, to incorporate a lateral smearing factor and range uncertainties (refer to the treatment planning section hereafter and the Appendix).

Treatment planning

Treatment planning was performed for 3DCRT, IMRT and PSPT at MAAS-TRO, the Netherlands Cancer Institute (NKI) and the Massachusetts General Hospital (MGH), using XiO (v4.34, CMS Software, St. Louis, MO), Pinnacle (ADAC Laboratories, Milpitas, CA) and a modified proton enabled version of XiO (v4.2.1), respectively. For the photon dose calculations, a multigrid superposition algorithm [28] or equivalent was used to account for tissue heterogeneity. Proton dose calculations were performed using a modified version of the XiO pencil beam algorithm with 1-dimensional density corrections. The dose

matrices, with a 3x3x3 mm^3 grid size, were uploaded in DICOM RT Dose format onto the secured central server.

For all modalities, the prescribed dose was 70 Gy to the PTV in 2 Gy fractions, once a day. Criteria for minimum and maximum dose were defined according to a modified ICRU 50 protocol [29]: $D_{98} \geq 95\%$ and $D_2 \leq 107\%$ of the prescribed dose. The maximum dose (D_2) to the spinal cord and oesophagus equalled 54 and 80 Gy respectively, using 2 Gy fractions, independent of the volume. For the lungs, the MLD was limiting to a biologically equivalent dose in 2 Gy fractions (EQD$_2$) of 19 Gy. The *MLD* volume was defined as the volume of both lungs minus the GTVs. For the heart, three DVH metrics were defined: V_{60} < 33%, V_{45} < 66% and V_{40} < 100%.

Two plans were created per patient. First, a fixed prescribed dose (PD) of 70 Gy to the tumour was used. Next, plans were up- or downscaled by adapting the fraction dose to fulfil all criteria of the OARs. The mean CTV dose was then reported as maximum tolerable dose (MTD). A maximum fraction dose of 4 Gy was considered, resulting in an EQD$_2$ to the tumour of 163 Gy. Table 2 summarises the physical and corresponding EQD$_2$ planning criteria for the target and OAR.

Table 2 Conversion table between EQD2 and physical dose for target and OAR

Structure of interest	α/β [Gy]	Criterion	EQD2 [Gy]	Physical dose at 35 fractions [Gy]
PTV	10	D$_2$	163	140
Lungs	3	MLD	≤19	Depends on distribution*
Heart	3	D33%	60	63
		D67%	45	51
		D100%	40	46
Oesophagus	3	D$_2$	80	77
Spinal cord	2	D$_2$	54	59

*The EQD2 of the MLD was calculated using the full 3D physical dose distribution and recalculating the equivalent dose at each point.

Photons

Three-dimensional conformal radiotherapy plans were created at MAASTRO conform the clinical guidelines. The plans consisted of multiple, optimised coplanar beams using a multi-leaf collimator (MLC) and wedges to shape the dose conformally to the target. In many cases, additional small beams from the same directions were used with limited dose to "pull" the isodoses around the PTV. The energy of the photon beams was almost always 10 MV.

Intensity-modulated radiotherapy plans were created at the NKI consisting of six to ten 10 MV beams. The beam configuration was mainly coplanar with a few exceptions. The collimator was rotated to fit best the shape of the target and OARs. For the optimisation, the criteria as given above were used in combination with constraints to some additional structures to control the dose to certain areas or force rapid dose fall-off around the PTV. Initial optimisation was done using 30 segments with 25 cm^2 size. When needed, the number of segments was increased to 50, while the segment size could be decreased to 12 cm^2.

Protons

The passive-scattered proton plans were planned at MGH using the XiO TPS. Each plan consisted of at least two (preferably three) beam directions to spread out the dose to normal tissues. Using a mid-ventilation CT-scan of the tumour, first, for each beam, an aperture, range compensator, range and modulation was chosen to conform the 95% isodose level as closely as possible to the target, *i.e.* the CTV. Second, margins were applied to the aperture, the range compensator and both the range and modulation width taking setup errors and breathing motion into account.

Beam directions that are parallel to density interfaces such as the lung and mediastinum were avoided because of the large range uncertainties even for small setup errors. Because of the low density of inflated lung (*e.g.* 0.25g/cm^3), overshoot due to range uncertainties and applied smearing results in a substantial volume of lung receiving full dose for a given beam direction. Therefore, the beam angles were intentionally aimed towards the mediastinum.

Refer to the Appendix for a detailed description of the proton planning process.

Analysis

To minimise the uncertainty in the analysis, all of the DVH metrics (refer to Table 3,4) were centrally derived by MAASTRO from the 3D dose matrices. Since the PTV is different for photon and proton modalities due to different margin algorithms, we chose the CTV for target comparison. We compared integral doses (ID), defined as the mean dose to the imaged patient, as dosimetric estimate for normal tissue toxicity differences between the three modalities.

For an estimation of the differences in the therapeutic windows of the treatment modalities, we introduced surrogate therapeutic indices (TIs) for the lungs, oesophagus and spinal cord. These were derived by dividing the mean

dose to the target by the essential dose metric of the structure. For example, for the lungs this would result in a TI_{MLD} defined by the MTD over the MLD.

To quantitatively assess the differences in conformity of the different treatment modalities, various indices have been proposed in the literature [30]. Because we wanted to compare the different treatment modalities based on target coverage and unwanted dose to the non-critical as well as the critical normal tissue (OARs), we used a "conformation number" (CN) as defined by:

$$CN = \frac{CTV_{95}}{CTV} \times \frac{CTV_{95}}{V_{95}} \qquad \text{(Eq. 1)}$$

where CTV_{95} and V_{95} are the volume of the CTV and the overall volume respectively, receiving minimally 95% of the target dose. Refer to the Appendix or [31] for a more elaborate description.

To investigate the volume of the low dose region in the patient, a sparing index ($\text{SPIN}_{50/10}$) was used [32]:

$$SPIN_{50/10} = 1 - V_{10}^{50} \qquad \text{(Eq. 2)}$$

where V_{10}^{50} is the ratio of the volume of tissue that receives between 10% and 50% of the prescribed dose to the irradiated volume. The irradiated volume was defined as the volume of tissue that receives at least 0.5% of the prescribed dose.

For both CN and $\text{SPIN}_{50/10}$, a value of 1 would indicate a perfect, theoretical dose distribution solely around the target and none in the healthy tissue.

Two-tailed, signed-rank Wilcoxon tests were calculated using SPSS (v15, Chicago, IL) and Matlab (The MathWorks, Natick, Massachusetts) to determine the significance of pair-wise differences between modalities. P-values less than 5/3% were considered significant.

We must note that, in general, some of the mentioned metrics are not suitable as absolute measures of plan quality or treatment modality performance. For instance, when considering volumes of OAR, they can depend on the imaged patient volume (*e.g.* spinal cord) and of course the anatomical differences of the patients. In this study, we used these metrics in a relative way and compared these metrics pair-wise as indicated above. This way, a statement of better plan quality is possible.

Results

Of the 25 included patients for the iso-effective protocol (Table 1), all were planned with 3DCRT and IMRT and 23 with PSPT. Two patients with tumours
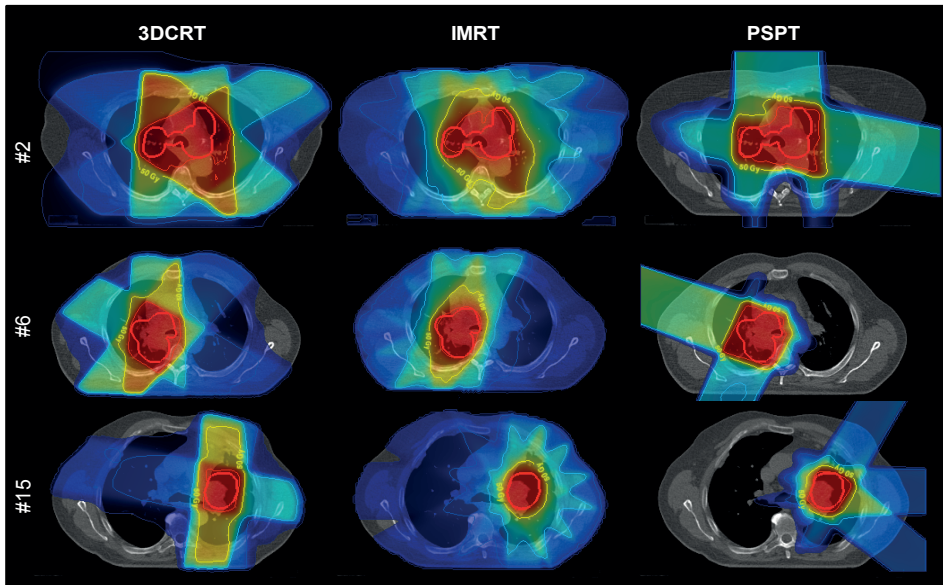
Figure 1 Comparison of dose distributions of 3DCRT, IMRT and PSPT treatment plans (columns) for three cases (rows). The target is shown (thick red line) with isodose lines of 1, 10, 25, 50 and 67 Gy (thin red line).
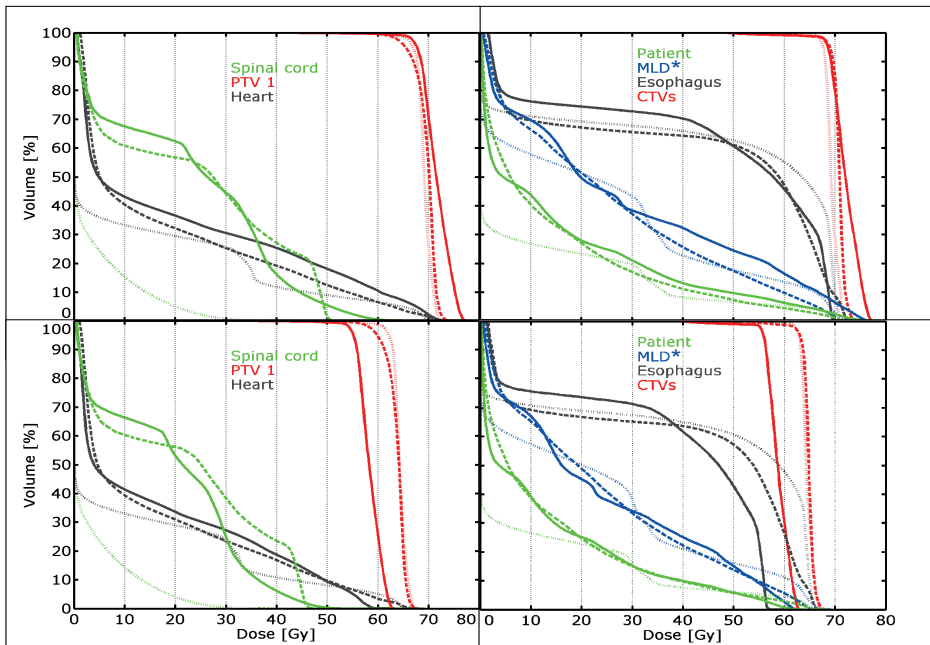
Figure 2 DVHs for patient #2 where the prescribed dose of 70 Gy resulted in exceeding the critical MLD level (upper row). Lowering the dose to the maximum tolerable dose resulted in lower DVH (bottom row), showing larger under-dosage for 3DCRT (solid) than for IMRT (dashed) or PSPT (dotted).

*MLD refers to the DVH of both left and right lung volumes minus the GTV.

located very cranially were excluded from proton planning. In real life, these patients would not be treated with PSPT because of the limited beam directions possible and because of remaining range uncertainties due to accuracy in the positioning of the arms. Because of the pair-wise analysis, we report on the remaining 23 corresponding datasets.

In Figure 1, mid-PTV dose distributions are shown for three typical cases. In general, IMRT showed a more conformal dose compared to 3DCRT and both showed a fairly large low dose region outside the target. PSPT showed its typical clear-edged beams with no exit dose. Figure 2 shows the DVHs for one of the patients (#2).

When irradiating with protons, it was found (Table 3) that, while prescribing the same dose to the target, the average ID was significantly lower than for 3DCRT (59%, $p < 0.001$) and IMRT (43%, $p < 0.001$). Except for V_{30Gy}, all lung volume metrics were significantly higher for photons than for protons. The average MLD was significantly higher for 3DCRT (40%; $p < 0.001$) and IMRT (21%; $p < 0.001$) when compared to PSPT. This resulted in a significantly higher TI_{MLD} for PSPT when compared to 3DCRT (51%, $p < 0.001$) or IMRT (39%, $p < 0.001$).

For the spinal cord and oesophagus, the average maximum doses (D_2) did not differ significantly between the three modalities. The average mean dose to the oesophagus was slightly larger for 3DCRT (16%, $p < 0.001$) and IMRT (6.6%, $p = 0.02$) than for PSPT. All average heart metrics were significantly lower for protons when compared to photons, except for the V_{65Gy}, which was 23% lower for IMRT ($p = 0.002$). The average TI_{Dmean} for both photon modalities was 44 times lower ($p < 0.001$) than for protons.

With respect to the conformity of the three different modalities, it was shown that IMRT had an average 50% higher CN ($p < 0.001$) while 3DCRT scored equally compared to PSPT. Also, the $SPIN_{50-10}$ was significantly worse for PSPT than for 3DCPT and IMRT: 0.61 ± 0.09 vs. 0.70 ± 0.05 ($p < 0.002$) and 0.72 ± 0.04 ($p < 0.001$), respectively.

Table 3 DVH metrics of the photon and proton treatment plans for the prescribed and maximum tolerable dose. Data for the target and the organs at risk are given as mean physical dose values with the standard deviation in parenthesis. Significant differences (p-values < 0.0167) of the photon versus the proton results are marked with an asterisk (*).

	PD = 70 Gy			MTD		
	Photons		Protons	Photons		Protons
	3DCRT	IMRT	PSPT	3DCRT	IMRT	PSPT
Target (CTV)						
D_{mean}	71.3 (1.0)*	70.7 (0.9)	70.3 (0.7)	75.0 (22.2)	81.5 (20.8)	75.0 (21.2)
D_2	73.9 (1.4)*	73.1 (1.9)*	72.0 (0.8)	77.6 (22.2)	84.2 (21.6)	76.7 (21.4)
D_{98}	68.9 (1.1)	68.7 (1.1)	68.7 (1.0)	72.6 (22.2)	79.2 (20.5)	73.3 (21.0)
CN	0.25 (0.07)	0.38 (0.10)*	0.25 (0.08)	0.25 (0.07)	0.38 (0.10)*	0.25 (0.08)
Patient						
ID	11.0 (5.4)*	9.9 (4.4)*	6.9 (3.9)	10.2 (3.2)*	10.6 (3.7)*	6.8 (3.1)
$SPIN_{50-10}$	0.70 (0.05)*	0.72 (0.04)*	0.61 (0.09)	0.71 (0.05)*	0.73 (0.04)*	0.62 (0.09)
TI_{ID}	8.6 (5.2)*	9.1 (5.4)*	15.3 (11.5)	8.6 (5.2)*	9.1 (5.4)*	15.3 (11.5)
Organs at risk						
Lung						
V_{30Gy}	21.0 (9.8)*	16.3 (7.7)	16.8 (8.9)	20.3 (7.6)*	17.7 (6.6)	16.7 (8.1)
V_{20Gy}	27.1 (12.6)*	23.4 (10.2)*	20.5 (10.4)	26.4 (9.5)*	25.0 (9.3)*	20.2 (10.2)
V_{13Gy}	37.0 (17.5)*	32.1 (11.8)*	23.3 (12.0)	36.1 (12.8)*	34.1 (10.2)*	23.1 (11.5)
V_{5Gy}	53.1 (17.0)*	56.9 (12.9)*	27.5 (13.9)	53.2 (15.5)*	59.1 (10.2)*	27.5 (13.7)
MLD	18.9 (7.3)*	16.4 (5.5)*	13.5 (6.2)	18.2 (4.1)*	17.9 (4.1)*	13.6 (5.6)
TI_{MLD}	4.5 (1.9)*	4.9 (2.0)*	6.8 (3.9)	4.5 (1.9)*	4.9 (2.0)*	6.8 (3.9)
Spinal cord						
D_2	40.0 (21.4)	42.6 (9.6)	37.9 (23.2)	37.0 (17.1)	46.8 (6.3)*	35.5 (18.1)
TI_{D2}	6.7 (15.2)	1.8 (0.7)*	31.7 (105)	6.8 (15.7)	1.8 (0.7)*	52.1 (198)
Oesophagus						
D_2	65.0 (15.7)	64.7 (15.8)	63.6 (17.9)	64.5 (13.8)	70.3 (11.0)	65.7 (17.3)
D_{mean}	28.3 (13.9)*	26.0 (12.1)	24.4 (13.7)	26.3 (9.4)*	27.5 (11.1)*	23.7 (11.7)
V_{55Gy}	31.0 (20.2)	26.4 (18.1)	28.3 (19.1)	22.0 (16.2)	27.3 (17.3)	22.2 (16.7)
V_{35Gy}	38.3 (22.7)*	34.9 (19.9)	35.3 (20.3)	37.8 (20.2)	37.0 (17.4)	35.3 (18.7)
TI_{D2}	1.3 (0.8)	1.3 (0.7)	2.0 (4.1)	1.3 (0.8)	1.3 (0.7)	2.0 (4.1)
Heart						
D_{mean}	15.3 (11.6)*	14.3 (10.3)*	7.6 (7.2)	14.3 (9.5)*	15.3 (10.0)*	7.5 (7.1)
V_{65Gy}	4.4 (5.5)*	2.3 (3.3)*	3.0 (4.0)	2.1 (3.0)	2.0 (2.9)	2.3 (4.2)
V_{45Gy}	13.3 (13.3)*	9.1 (9.3)*	6.2 (6.5)	10.1 (9.6)*	9.6 (8.5)*	6.1 (6.3)
V_{40Gy}	15.1 (14.5)*	11.5 (11.7)*	7.2 (7.2)	12.4 (10.8)*	12.4 (10.9)*	7.2 (7.1)
V_{30Gy}	19.5 (17.2)*	17.0 (16.1)*	12.0 (13.4)	18.5 (15.1)*	18.5 (15.4)*	11.7 (13.1)
V_{20Gy}	27.0 (23.7)*	25.3 (22.2)*	15.1 (15.6)	26.5 (19.8)*	27.6 (21.3)*	15.0 (15.4)
V_{10Gy}	40.9 (31.8)*	41.0 (31.9)*	18.1 (17.5)	40.7 (30.6)*	45.0 (33.4)*	18.1 (17.3)
TI_{Dmean}	13.1 (16.8)*	13.0 (14.9)*	572 (2074)	13.1 (16.8)*	13.0 (14.9)*	572 (2074)

Abbreviations: PD = prescribed dose; MTD = maximum tolerable dose; 3DCRT / IMRT / PSPT = 3D conformal photon / intensity-modulated photon / 3D passive scattered proton radiotherapy; Dx = dose (Gy) given to x% (or mean) of the volume; CN = conformation number (Eq. 4); ID = integral dose in Gy; TIy = therapeutic index (MTD / y); VzGy = percent volume of the total organ's volume that receives more than z Gy (for the lung results, the volume was taken as both lungs minus the PTV for photons); MLD = mean lung dose in Gy (lung volume taken as both lungs minus the GTVs)

Table 4 DVH metrics of the photon and proton treatment plans, split for plans where all modalities simultaneously could escalate the dose to the tumour and some that did not. Metric notification identical to Table 3.

	One or more modalities had to de-escalate (n=13)			All modalities could escalate (n=10)		
	Photons		Protons	Photons		Protons
	3DCRT	IMRT	PSPT	3DCRT	IMRT	PSPT
Target (CTV)						
D_{mean}	66.7 (22.8)	76.6 (19.2)	64.6 (10.1)	85.8 (17.0)	87.9 (22.0)	88.5 (24.5)
D_2	69.4 (22.7)	79.0 (19.6)	66.2 (10.5)	88.3 (17.3)	91.1 (23.2)	90.4 (24.6)
D_{98}	64.3 (22.8)	74.4 (19.0)	63.1 (10.0)	83.5 (16.7)	85.6 (21.7)	86.4 (24.5)
CN	0.26 (0.08)	0.41 (0.12)*	0.28 (0.10)	0.23 (0.06)	0.34 (0.07)*	0.23 (0.05)
Patient						
ID	11.8 (3.1)*	12.3 (3.7)*	8.1 (3.1)	8.2 (2.1)*	8.3 (2.1)*	5.0 (2.0)
$SPIN_{50-10}$	0.69 (0.05)*	0.72 (0.04)*	0.63 (0.08)	0.74 (0.04)*	0.74 (0.03)*	0.60 (0.10)
TI_{ID}	6.7 (5.4)*	7.4 (5.4)*	11.5 (11.6)	11.1 (3.9)*	11.4 (4.6)*	20.2 (9.9)
Organs at risk						
Lung						
V_{30Gy}	22.6 (8.4)	18.8 (6.8)	18.8 (9.2)	17.2 (5.4)	16.4 (6.4)*	13.9 (5.9)
V_{20Gy}	30.1 (9.9)*	26.9 (9.7)	23.4 (11.2)	21.6 (6.7)*	22.5 (8.5)*	16.1 (7.2)
V_{13Gy}	40.2 (14.5)*	36.5 (10.3)*	27.1 (12.5)	30.8 (7.9)*	31.0 (9.6)*	17.8 (7.8)
V_{5Gy}	60.4 (15.7)*	63.7 (7.6)*	32.3 (15.1)	43.9 (9.4)*	53.2 (10.5)*	21.4 (9.0)
MLD	19.8 (3.6)*	19.5 (3.3)*	15.4 (5.7)	16.1 (3.8)*	15.8 (4.2)*	11.3 (4.8)
TI_{MLD}	3.6 (1.8)*	4.2 (1.9)*	4.9 (2.5)	5.6 (1.5)*	5.9 (1.9)*	9.2 (4.2)
Spinal cord						
D_2	42.3 (16.2)	48.3 (4.8)	44.7 (13.2)	30.1 (16.5)	44.8 (7.7)*	23.6 (17.1)
TI_{D2}	6.9 (19.5)	1.6 (0.37)	1.7 (0.87)	6.7 (9.5)	2.1 (0.9)*	117 (295)
Oesophagus						
D_2	59.0 (9.5)	69.8 (11.5)	61.2 (19.8)	71.7 (15.6)	71.0 (10.7)	71.5 (11.9)
D_{mean}	31.2 (7.6)	33.1 (10.7)	29.8 (10.7)	20.0 (7.9)*	20.2 (6.6)*	15.9 (7.7)
V_{55Gy}	25.5 (18.5)	35.1 (17.1)	26.8 (19.0)	17.6 (12.0)	17.2 (11.8)	16.3 (11.6)
V_{35Gy}	49.5 (16.3)	45.5 (17.7)	45.5 (17.2)	22.5 (13.5)	26.0 (9.1)*	22.0 (10.6)
TI_{D2}	1.2 (0.78)	1.2 (0.81)	2.5 (5.4)	1.3 (0.8)	1.3 (0.59)	1.3 (0.72)
Heart						
D_{mean}	17.3 (9.3)*	18.9 (9.8)*	9.6 (7.9)	10.4 (8.7)*	10.5 (8.5)*	4.7 (5.0)
V_{65Gy}	1.3 (2.5)	2.2 (3.6)	2.5 (5.4)	3.1 (3.5)	1.7 (1.7)	2.0 (2.1)
V_{45Gy}	10.9 (9.1)	12.5 (9.0)*	7.8 (7.3)	9.0 (10.6)*	5.7 (6.1)	4.0 (4.0)
V_{40Gy}	14.3 (10.1)	16.3 (11.5)*	9.2 (8.2)	10.0 (11.6)*	7.2 (7.9)*	4.6 (4.6)
V_{30Gy}	23.3 (14.9)	24.1 (15.6)*	15.5 (15.3)	12.2 (13.4)*	11.3 (12.2)*	6.7 (7.7)
V_{20Gy}	33.2 (19.2)	35.0 (21.0)*	19.8 (17.2)	17.8 (17.8)*	17.9 (18.3)*	8.8 (10.5)
V_{10Gy}	53.2 (31.8)*	56.1 (31.6)*	23.6 (19.0)	24.4 (20.7)*	30.5 (31.4)*	10.9 (12.1)
TI_{Dmean}	11.3 (19.8)*	9.0 (12.6)*	941 (2744)	15.4 (12.3)*	18.0 (16.8)*	91.5 (126.8)

Abbreviations identical to Table 3.

After up- or downscaling all plans to clinically acceptable MTD by fulfilling all planning criteria, the prescribed dose to the tumour of 70 Gy could be increased for 52%, 83% and 61% of the patients for 3DCRT, IMRT and PSPT, respectively (Table 3). The other patients received the prescribed dose or lower. Refer to Figure 2 for DVHs of such a case (#2).

The mean MTD was 8.6% higher for IMRT when compared to PSPT, although this was not significant. The 3DCRT MTD did not differ from PSPT. The integral dose was higher again for both photons modalities when compared to protons (p=0.005). Once more, all lung volume metrics were significantly higher for photons than for protons, except for V_{30Gy}. The average MLD was significantly higher for 3DCRT (34%; p<0.001) and IMRT (32%; p<0.001) when compared to PSPT.

For the spinal cord, the average maximum dose (D_2) was 42% higher for IMRT than for PSPT (p=0.005). The average mean dose to the oesophagus was larger for 3DCRT (11%, p=0.007) and IMRT (16%, p=0.005) than for PSPT. All average heart metrics were significantly lower for protons when compared to photons, except for the V_{65Gy}, which did not differ significantly.

Due to the rescaling of the dose matrices, the DVH metrics that are based on relative dose levels or ratios typically do not change. One exception is the TI_{D2} for the spinal cord where the accuracy of the D_2 is limited by some very small irradiated volumes.

For 10 patients (two stage I, one stage II and 7 stage III), all modalities could escalate the dose to an MTD of around 87 Gy (Table 4). The average integral dose was about 65% higher for photons than for protons (p=0.005). The MLD was 42% higher for 3DCRT (p=0.007) than for PSPT, while it was 40% higher for IMRT (p=0.005). The TI_{ID} and TI_{MLD} were nearly double of those of the non-escalated group. For the spinal cord, the maximum dose was 90% higher for IMRT than for PSPT (p=0.005). For the oesophagus, the maximum dose was approximately 71.5 Gy for all modalities. The mean dose to the heart for photons was double the dose for protons, while the V_{40Gy} was 117% higher for 3DCRT (p=0.013) and 57% higher for IMRT (p=0.013) than for PSPT.

Further sub-group analysis based on tumour volume, location or stage did not provide new information and is omitted from this report.

Discussion

With the first results from 23 cases of the lung cohort of the ROCOCO trial, we have shown that for the treatment of NSCLC, Stage I to III, proton radiotherapy is superior to photon radiotherapy, in terms of reduction of the normal tissue dose. Based on the integral dose, an absolute improvement of 3-4 Gy was seen for protons, resulting in an improved therapeutic index in the order of 15 compared to 9.

In the current literature, with a few exceptions [33-35], ten or less patients were used in the published treatment planning comparisons. Furthermore, in most of the published papers it is unclear what the selection criteria were of the included patients, possibly introducing selection bias. Publications dating before 2000 did not consider intensity modulated radiation therapy as a modality for photon therapy or intensity-modulated particle therapy. By using a multicentric approach, with consensus on a predefined protocol and a relatively large amount of patients, a high level of confidence was achieved in the current study as the established clinical experience of the participants was used.

In the present study we found that the average V_{20Gy} of the lungs were lower for both IMRT and PSPT compared to 3DCRT (16.3%, 16.8% and 21%, respectively). Regarding the MLD, it was found that this was lowest for PSPT (13.5 Gy) when compared to 3DCRT (18.9 Gy) or IMRT (16.4 Gy). Previous research has shown that the V_5 of the lungs was also associated with lung toxicity [36]. This study showed that the V_5 for protons was less than half the value for both photon modalities.

While the average maximum dose to the spinal cord did not significantly change between all three modalities, it was shown that the average of the corresponding individual therapeutic index was significantly higher for PSPT (31.7) than for IMRT (1.8). This indicates that there are a few outliers present in the derived metrics.

Furthermore, the maximum dose to the oesophagus was nearly equal for all three modalities. The V_{55Gy} appeared to be lowest for IMRT (26.4%). When considering the V_{35Gy} however, it was equal for IMRT and PSPT (35%). The mean heart dose for photons was nearly double the dose of protons, while the V_{40Gy}, which is known for its correlation to heart toxicity was 110% and 60% higher for 3DCRT and IMRT than for PSPT [24].

It is expected that the lower dose to the normal tissues with proton therapy will lower the probability of normal tissue complications and result in a better quality of life in an iso-effective setting. Clinical validated models have been published showing this for a variety of acute and late toxicities [37-40].

Radiation therapy is known for its statistically significant increased risk of secondary malignancies [41]. The hypothesis of reduced late side effects with proton therapy has been investigated intensely and tends to be correctly supported, although large uncertainties in the relative biological effect, among others, remain [42-46]. Additionally, improvements in the treatment delivery equipment further reduce the amount of secondary neutron production, which is the main contributor to the scattered dose to the normal tissue [47-49]. Although the occurrence of secondary tumours for lung cancer patients could be considered irrelevant due to the short life expectancy, the expected reduced carcinogenic risk should be mentioned with respect to the changing NSCLC population and overall treatment improvements.

Another advantage of the small low dose volume and limited number of beams for protons is that it offers the opportunity to better re-irradiate for loco-regional metastasis.

The increased therapeutic ratios showed that there is room for iso-toxic dose-escalation with protons for certain patients. A unicentric study including a selection of 15 stage III patients has shown that dose-escalation from 63 to 74 Gy was possible with PSPT while keeping normal tissue toxicity lower than with 3DCRT or IMRT [33]. Early results of a subsequent phase II trial show that a dose prescription to 74 Gy is well-tolerated with proton therapy [50].

The current report predicts that this still holds for MTDs exceeding these levels with about 10 Gy. For 40% of the patients, dose-escalation was possible to a mean level of more than 85 Gy for all modalities. With dose levels of such ablative magnitude, while still maintaining the dose to the OAR below a toxic level, it is expected that local tumour control will significantly be increased and consequently will lead to a further improved survival with acceptable toxicity. A phase II dose-escalation trial is currently being performed to investigate sub-boosting of high FDG-PET uptake regions to such high doses with IMRT [51]. PSPT will not be part of this trial because range uncertainties make it nearly impossible to ensure a certain limited dose to an OAR by means of forward planning.

When the treatment plans of the three modalities were evaluated for possible dose-escalation, it showed that sometimes the dose could be increased for one modality while it had to be lowered for another. The fact that this heterogeneity

cannot be predicted beforehand makes generalisation of the achieved results difficult. This raises the demand for an individualised approach when trying to classify eligibility of a patient to receive proton therapy instead of less expensive photon therapy (see Slides, Supplemental Digital Content 1, which show a Proton therapy reimbursement decision tree for the Netherlands and figures 1 and 2).

The performance of the CN and SPIN$_{50-10}$ is worse when using PSPT instead of 3DCRT or IMRT. The error margins that are needed to compensate in depth and width and the fact that only two to three beams are used generate an irradiated volume that exceeds the PTV used for photons. However, even though PSPT suffers from these disadvantages, the dose to the normal tissues remains much lower than for 3DCRT and IMRT. This indicates that the use of conformity and dose spill indices is useful for plan comparison but they do not cover the full plan evaluation and cannot by themselves indicate prevalence of either treatment modality.

One might argue that passive scattered proton therapy is not the latest and most advanced type of particle therapy, but in spite of this the current results show that it is favourable in terms of normal tissue dose reduction when compared to photon radiation therapy. It could be argued that if results were derived from modern scanning proton beam therapy, the dose distributions could improve due to reduced secondary neutron production [47,52] and improved conformity. Recent publications show indeed that intensity-modulated proton therapy (IMPT) is able to lower the dose to the normal tissue and allows dose-escalation up to 88 Gy for a selection (n=20) of extensive stage IIIB NSCLC patients [53].

For stage I NSCLC patients, stereotactic radiotherapy (SRT) is increasingly used. In a meta-analysis, it was previously found that proton therapy results in similar survival rates as SBRT [9]. In another publication, the influence of different breathing suppression methods was investigated in a PSPT, IMPT and SRT treatment planning study [34]. Both proton techniques were superior in sparing the normal tissue when compared to SBRT. However, the differences were small and improved local control rates needed clinical validation. More recent publications report that PSPT and especially IMPT do benefit from the superior dose distribution and result in significant normal tissue sparing [54,55].

Within our current dataset, there were individual cases (stage IIA and IIIA) that reached MTD levels well above 100 Gy with any of the three modalities. Subgroup analysis did not show a significantly higher proton dose level for tumours with different staging or tumour size. This again indicates that the best treatment

modality could be difficult to predict beforehand and should be investigated on an individual basis. Future research should preferably include other modalities such as IMPT or even C-ion radiotherapy. The ROCOCO consortium is considering to do so with the datasets and methods presented here. However, up to this date we have not found a reliable method to calculate treatment plans for these new treatment modalities when delivered to moving lung tumours.

The evidence for the dosimetric improvements of particle therapy is still hampered by sufficient clinically validated results [13-19,56]. It is however possible to use highly accurate dose calculations and well-established predictive radiobiological models [57-59] in an *in silico* approach as a surrogate to determine the (cost-)effectiveness of particle therapy with sufficient reliability. A signal in that direction is the fact that the *in silico* approach used in this study was acknowledged and justified by the Dutch Health Care Insurance Board (CVZ) [60] as a vital supplement to prospective RCTs.

Recently published studies show the potential cost-effectiveness of particle therapy in NSCLC. However, they emphasise the uncertainty in determining this and the probability of making a wrong decision with regard to establishing a particle centre [61]. It is, therefore, important that when new evidence becomes available reassessment of the (cost-)effectiveness of particle therapy in lung cancer should be carried out. Next, the theoretical benefit should be confirmed by clinical evidence from well-designed prospective studies. We argue that the results of *in silico* trials such as the one currently presented aid to choose the most relevant areas of research for RCTs involving proton therapy.

Conclusions

Using an *in silico* approach, we found that while maintaining a good coverage of the target, proton radiotherapy significantly reduced the dose to the normal tissue, as indicated by a lower integral dose, when compared to conventional or intensity-modulated photon therapy of NSCLC patients.

Furthermore, the presented data show that dose escalation is possible and that an increased local tumour control can be expected, hence improving survival. We believe that carefully designed RCTs should now be performed to validate these results.

The current findings provide an incentive to investigate other tumour sites and modalities such as C-ion radiotherapy in an *in silico* set-up. For NSCLC patients, the possibilities of hypo-fractionated delivery schedules could be inves-

tigated as well. Finally, since investment costs are high, it would be valuable to investigate whether particle therapy can be considered cost-effective.

Acknowledgments

Appendix

Margin recipe

The clinical target volume (CTV) was defined as the GTV with a margin of 5 mm. Next, the individual, anisotropic planning target volume (PTV) margins for photons (Eq. A1) were calculated in three directions (CC, LR and AP) using a margin recipe [62]

$$M_{PTV} = 2.5 \cdot \Sigma + 0.7\sqrt{\sigma^2 + A^2/8}$$

(Eq. A1)

with Σ the overall standard deviation (SD) of the systematic errors, σ the overall SD of the random errors and A the peak-to-peak amplitude of the tumour. We used the systematic and random set-up errors as determined in MAASTRO's clinical setting. Because the mid-ventilation CT was used there was no systematic motion error included.

Proton Treatment Planning

The passive-scattered proton plans were planned using the XiO TPS (v4.2.1, CMS Software, St. Louis, MO). In passive-scattered proton radiotherapy, each beam delivers a homogenous dose to the target volume. In order to spread-out the dose to normal tissues each plan consists of at least two, but preferably three, beam directions. Although range uncertainties and setup errors are taken into account in the treatment planning process, the use of multiple beam directions minimizes the risk of underdosing the target due to unexpected density variations due to *e.g.* heavy breathing or a substantial change in the average tumour position with respect to the patient anatomy as used for treatment planning.

Given a mid-ventilation CT-scan of the tumour and a description of the breathing-induced excursion from this position in all three directions, treatment planning of proton radiotherapy at MGH is a two-step process. As we may choose to treat only a subset of fields on any given treatment day, it is important that each field separately ensures target coverage.

First, for each beam, we conform the 95% isodose level as closely as possible to the target, *i.e.* the CTV. The aperture shape for each beam is chosen to conform the 95% isodose level in all lateral directions. The proton beam range and range compensator are chosen to conform distally (downstream) to the target volume. The radiological "thickness" of the target (in the depth direction) determines the choice of modulation width of the spread-out Bragg peak (SOBP). In our centre, we employ M_{98} for the definition of modulation, *i.e.* the distance between the proximal 98% and the distal 90% isodose level. Because of the nature of passive-scattered proton radiotherapy, tight proximal coverage cannot be achieved except for those regions where the radiological thickness equals the maximum thickness.

Second, margins were applied to the aperture, the range compensator and both the range and modulation width. Range uncertainties of 3.5% and 1 mm were applied [63]. The same range uncertainty to the modulation width was applied, taking into account that the increase in range already translated into an increase in necessary modulation width. For example, a field with a range of 16 cm and modulation of 10 cm after step one, will have a range and modulation width of 16.8 cm and 11.2 cm, respectively.

Aperture expansion and range compensator smearing was applied [64], meaning that the safety margin is typically less than the summation of setup error and half the peak-to-peak breathing amplitude. Aperture expansion (*i.e.* lateral margining) compensates for setup errors of the tumour in the lateral direction with respect to the central beam axis. Range compensator smearing (*i.e.* distal margining) is applied to take into account the effect of the shift in the patient density distribution and the detrimental effects of these density variations on distal target coverage. The overshoot due to smearing depends on the local variation in the range compensator thickness and hence is not uniform across the lateral extent of a field. The exact magnitude of lateral and distal margining depends on the extent of the breathing motion, but will never be less than the expected maximum setup error, refer to Table A1.

Our TPS only allows a single uniform value for smearing on a per range compensator basis. Lateral aperture margins differed depending on the breath-

ing motion in a specific direction. All patients' plans consisted of fields in the transversal plane only.

Conformity index

Assessment of the differences in conformity when comparing treatment plans from different modalities has been published before [30]. To compare target coverage and unwanted dose to the non-critical as well as the critical normal tissue (OARs), we chose to use the "conformation number" (CN) [31].

The CN takes into account the quality of tumour irradiation as indicated by the first part of the equation (Eq. A2) and the irradiation of the non-critical tissue, indicated by the second part.

$$CN_{T,ref} = \frac{V_{T,ref}}{V_T} \times \frac{V_{T,ref}}{V_{ref}} \qquad \text{(Eq. A2)}$$

where $V_{T,ref}$ is the volume of the target receiving a dose equal to or greater than the reference dose, V_T is the volume of the target, V_{ref} is the overall volume receiving a dose equal to or greater than the reference dose or target dose [29]. The reference dose was chosen to be 95% of the prescribed dose to the PTV and the CTV was again chosen as the target volume. For clarity, the CN can thus be written as follows in our case:

$$CN = \frac{CTV_{95}}{CTV} \times \frac{CTV_{95}}{V_{95}} \qquad \text{(Eq. A2)}$$

References

[1] Jemal A, Siegel R, Ward E, Murray T, Xu J Thun MJ. Cancer statistics, 2007. CA Cancer J Clin 2007;57:43-66.

[2] Ferlay J, Autier P, Boniol M, Heanue M, Colombet M Boyle P. Estimates of the cancer incidence and mortality in Europe in 2006. Ann Oncol 2007;18:581-592.

[3] De Ruysscher D, Dehing C, Bentzen SM, et al. Can we optimize chemo-radiation and surgery in locally advanced stage III non-small cell lung cancer based on evidence from randomized clinical trials? A hypothesis-generating study. Radiother Oncol 2009;93:389-395.

[4] Brada M, Pijls-Johannesma M De Ruysscher D. Proton therapy in clinical practice: current clinical evidence. J Clin Oncol 2007;25:965-970.

[5] Brada M, Pijls-Johannesma M De Ruysscher D. Current clinical evidence for proton therapy. Cancer J 2009;15:319-324.

[6] Lodge M, Pijls-Johannesma M, Stirk L, Munro AJ, De Ruysscher D Jefferson T. A systematic literature review of the clinical and cost-effectiveness of hadron therapy in cancer. Radiother Oncol 2007;83:110-122.

[7] Olsen DR, Bruland OS, Frykholm G Norderhaug IN. Proton therapy - a systematic review of clinical effectiveness. Radiother Oncol 2007;83:123-132.

[8] Pijls-Johannesma M, Grutters JP, Lambin P Ruysscher DD. Particle therapy in lung cancer: where do we stand? Cancer Treat Rev 2008;34:259-267.

[9] Grutters JP, Kessels AG, Pijls-Johannesma M, De Ruysscher D, Joore MA Lambin P. Comparison of the effectiveness of radiotherapy with photons, protons and carbon-ions for non-small cell lung cancer: a meta-analysis. Radiother Oncol 2009;95:32-40.

[10] Pijls-Johannesma M, Grutters JP, Verhaegen F, Lambin P De Ruysscher D. Do we have enough evidence to implement particle therapy as standard treatment in lung cancer? A systematic literature review. Oncologist 2010;15:93-103.

[11] Pijls-Johannesma M, Pommier P Lievens Y. Cost-effectiveness of particle therapy: current evidence and future needs. Radiother Oncol 2008;89:127-134.

[12] Zietman A, Goitein M Tepper JE. Technology evolution: is it survival of the fittest? J Clin Oncol 2010;28:4275-4279.

[13] Bentzen SM. Randomized controlled trials in health technology assessment: overkill or overdue? Radiother Oncol 2008;86:142-147.

[14] Glatstein E, Glick J, Kaiser L Hahn SM. Should randomized clinical trials be required for proton radiotherapy? An alternative view. J Clin Oncol 2008;26:2438-2439.

[15] Glimelius B Montelius A. Proton beam therapy - do we need the randomised trials and can we do them? Radiother Oncol 2007;83:105-109.

[16] Goitein M Cox JD. Should randomized clinical trials be required for proton radiotherapy? J Clin Oncol 2008;26:175-176.

[17] Tepper JE. Protons and parachutes. J Clin Oncol 2008;26:2436-2437.

[18] Introducing new treatments and technologies into the clinic. J Clin Oncol 2008;26:2425.

[19] Suit H, Kooy H, Trofimov A, et al. Should positive phase III clinical trial data be required before proton beam therapy is more widely adopted? No. Radiother Oncol 2008;86:148-153.

[20] Roelofs E, Persoon L, Qamhiyeh S, et al. Design of and technical challenges involved in a framework for multicentric radiotherapy treatment planning studies. Radiother Oncol 2010;97:567-571.

[21] van Baardwijk A, Wanders S, Boersma L, et al. Mature results of an individualized radiation dose prescription study based on normal tissue constraints in stages I to III non-small-cell lung cancer. J Clin Oncol 2010;28:1380-1386.

[22] Bosmans G, Buijsen J, Dekker A, et al. An "in silico" clinical trial comparing free breathing, slow and respiration correlated computed tomography in lung cancer patients. Radiother Oncol 2006;81:73-80.

[23] Chapet O, Kong FM, Quint LE, et al. CT-based definition of thoracic lymph node stations: an atlas from the University of Michigan. Int J Radiat Oncol Biol Phys 2005;63:170-178.

[24] De Ruysscher D, Faivre-Finn C, Nestle U, et al. European Organisation for Research and Treatment of Cancer recommendations for planning and delivery of high-dose, high-precision radiotherapy for lung cancer. J Clin Oncol 2010;28:5301-5310.

[25] Daisne JF, Sibomana M, Bol A, Doumont T, Lonneux M Gregoire V. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. Radiother Oncol 2003;69:247-250.

[26] Ollers M, Bosmans G, van Baardwijk A, et al. The integration of PET-CT scans from different hospitals into radiotherapy treatment planning. Radiother Oncol 2008;87:142-146.

[27] van Baardwijk A, Bosmans G, Bentzen SM, et al. Radiation dose prescription for non-small-cell lung cancer according to normal tissue dose constraints: an in silico clinical trial. Int J Radiat Oncol Biol Phys 2008;71:1103-1110.

[28] Miften M, Wiesmeyer M, Monthofer S Krippner K. Implementation of FFT convolution and multigrid superposition models in the FOCUS RTP system. Phys Med Biol 2000;45:817-833.

[29] ICRU, International Commission on radation Units and Measurements report 50. Prescribing, recording and reporting photon beam therapy, 1993.

[30] Feuvret L, Noel G, Mazeron JJ Bey P. Conformity index: a review. Int J Radiat Oncol Biol Phys 2006;64:333-342.

[31] van't Riet A, Mak AC, Moerland MA, Elders LH van der Zee W. A conformation number to quantify the degree of conformality in brachytherapy and external beam irradiation: application to the prostate. Int J Radiat Oncol Biol Phys 1997;37:731-736.

[32] Al-Yahya K, Schwartz M, Shenouda G, Verhaegen F, Freeman C Seuntjens J. Energy modulated electron therapy using a few leaf electron collimator in combination with IMRT and 3D-CRT: Monte Carlo-based planning and dosimetric evaluation. Med Phys 2005;32:2976-2986.

[33] Chang JY, Zhang X, Wang X, et al. Significant reduction of normal tissue dose by proton radiotherapy compared with three-dimensional conformal or intensity-modulated radiation therapy in Stage I or Stage III non-small-cell lung cancer. Int J Radiat Oncol Biol Phys 2006;65:1087-1096.

[34] Georg D, Hillbrand M, Stock M, Dieckmann K Pötter R. Can protons improve SBRT for lung lesions? Dosimetric considerations. Radiother Oncol 2008;88:368-375.

[35] Wang C, Nakayama H, Sugahara S, Sakae T Tokuuye K. Comparisons of dose-volume histograms for proton-beam versus 3-D conformal x-ray therapy in patients with stage I non-small cell lung cancer. Strahlenther Onkol 2009;185:231-234.

[36] Wang S, Liao Z, Wei X, et al. Analysis of clinical and dosimetric factors associated with treatment-related pneumonitis (TRP) in patients with non-small-cell lung cancer (NSCLC) treated with concurrent chemotherapy and three-dimensional conformal radiotherapy (3D-CRT). Int J Radiat Oncol Biol Phys 2006;66:1399-1407.

[37] De Ruysscher D, Dehing C, Bremer RH, et al. Maximal neutropenia during chemotherapy and radiotherapy is significantly associated with the development of acute radiation-induced dysphagia in lung cancer patients. Ann Oncol 2007;18:909-916.

[38] Dehing-Oberije C, De Ruysscher D, Petit S, et al. Development, external validation and clinical usefulness of a practical prediction model for radiation-induced dysphagia in lung cancer patients. Radiother Oncol 2010;97:455-461.

[39] Marks LB, Bentzen SM, Deasy JO, et al. Radiation dose-volume effects in the lung. Int J Radiat Oncol Biol Phys 2010;76:S70-76.

[40] Werner-Wasik M, Yorke E, Deasy J, Nam J Marks LB. Radiation dose-volume effects in the esophagus. Int J Radiat Oncol Biol Phys 2010;76:S86-93.

[41] ICRU, International Commission on radation Units and Measurements report 83. Prescribing, Recording, and Reporting Photon-Beam Intensity-Modulated Radiation Therapy (IMRT), 2010.

[42] Paganetti H. Nuclear interactions in proton therapy: dose and relative biological effect distributions originating from primary and secondary particles. Phys Med Biol 2002;47:747-764.

[43] Blakely EA Chang PY. Late effects from hadron therapy. Radiother Oncol 2004;73 Suppl 2:S134-140.

[44] Hall EJ. Intensity-modulated radiation therapy, protons, and the risk of second cancers. Int J Radiat Oncol Biol Phys 2006;65:1-7.

[45] Brenner DJ Hall EJ. Secondary neutrons in clinical proton radiotherapy: a charged issue. Radiother Oncol 2008;86:165-170.

[46] Kim S, Min BJ, Yoon M, et al. Secondary radiation doses of intensity-modulated radiotherapy and proton beam therapy in patients with lung and liver cancer. Radiother Oncol 2011;98:335-339.

[47] Schneider U, Agosteo S, Pedroni E Besserer J. Secondary neutron dose during proton therapy using spot scanning. Int J Radiat Oncol Biol Phys 2002;53:244-251.

[48] Taddei PJ, Fontenot JD, Zheng Y, et al. Reducing stray radiation dose to patients receiving passively scattered proton radiotherapy for prostate cancer. Phys Med Biol 2008;53:2131-2147.

[49] Brenner DJ, Elliston CD, Hall EJ Paganetti H. Reduction of the secondary neutron dose in passively scattered proton radiotherapy, using an optimized pre-collimator/collimator. Phys Med Biol 2009;54:6065-6078.

[50] Chang JY, Komaki R, Lu C, et al. Phase 2 study of high-dose proton therapy with concurrent chemotherapy for unresectable stage III nonsmall cell lung cancer. Cancer; 2011/03/26; [Epub ahead of print]

[51] De Ruysscher D. Dose Escalation by Boosting Radiation Dose in Stage II and III Non Small Cell Lung Cancer (NSCLC): a Phase II Trial (BTV Boost);Accessed 2010-04-23, at clinicaltrials.gov/ct2/show/NCT01024829.

[52] Shin D, Yoon M, Kwak J, et al. Secondary neutron doses for several beam configurations for proton therapy. Int J Radiat Oncol Biol Phys 2009;74:260-265.

[53] Zhang X, Li Y, Pan X, et al. Intensity-modulated proton therapy reduces the dose to normal tissue compared with intensity-modulated radiation therapy or passive scattering proton therapy and enables individualized radical radiotherapy for extensive stage IIIB non-small-cell lung cancer: a virtual clinical study. Int J Radiat Oncol Biol Phys 2009;77:357-366.

[54] Kadoya N, Obata Y, Kato T, et al. Dose-volume comparison of proton radiotherapy and stereotactic body radiotherapy for non-small-cell lung cancer. Int J Radiat Oncol Biol Phys 2010;79:1225-1231.

[55] Register SP, Zhang X, Mohan R Chang JY. Proton stereotactic body radiation therapy for clinically challenging cases of centrally and superiorly located stage I non-small-cell lung cancer. Int J Radiat Oncol Biol Phys 2010;80:1015-1022.

[56] Glimelius B, Ask A, Bjelkengren G, et al. Number of patients potentially eligible for proton therapy. Acta Oncol 2005;44:836-849.

[57] The 2007 Recommendations of the International Commission on Radiological Protection. ICRP publication 103. Ann ICRP 2007;37:1-332.

[58] Bentzen SM, Constine LS, Deasy JO, et al. Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC): an introduction to the scientific issues. Int J Radiat Oncol Biol Phys 2010;76:S3-9.

[59] Marks LB, Yorke ED, Jackson A, et al. Use of normal tissue complication probability models in the clinic. Int J Radiat Oncol Biol Phys 2010;76:S10-19.

[60] Ligtenberg G, Staal P Kleijnen S, Rapport Protonentherapie [Proton therapy]. Dutch Health Care Insurance Board (CVZ) 2009.

[61] Grutters JP, Pijls-Johannesma M, Ruysscher DD, et al. The cost-effectiveness of particle therapy in non-small cell lung cancer: exploring decision uncertainty and areas for future research. Cancer Treat Rev 2010;36:468-476.

[62] van Herk M. Errors and margins in radiotherapy. Semin Radiat Oncol 2004;14:52-64.

[63] Moyers MF, Miller DW, Bush DA Slater JD. Methodologies and tools for proton beam design for lung tumors. Int J Radiat Oncol Biol Phys 2001;49:1429-1438.

[64] Engelsman M Kooy HM. Target volume dose considerations in proton beam treatment planning for lung tumors. Med Phys 2005;32:3549-3557.

Chapter 7

PRODECIS: Proton Decision Support

Development and Evaluation of an Online Three-Level Proton vs Photon Decision Support Prototype for Head and Neck Cancer – Comparison of Dose, Toxicity and Cost-Effectiveness

Qing Cheng*, Erik Roelofs*, Bram L.T. Ramaekers, Daniëlle Eekers, Johan van Soest, Tim Lustberg, Tim Hendriks, Frank Hoebers, Hans Paul van der Laan, Erik W. Korevaar, Andre Dekker, Johannes A. Langendijk, Philippe Lambin; *Equal contribution

Abstract

To quantitatively assess the effectiveness of proton therapy for individual patients, we developed a prototype for an online platform for proton decision support (PRODECIS) comparing photon and proton treatments on dose metric, toxicity and cost-effectiveness levels. An evaluation was performed with 23 head and neck cancer datasets.

Introduction

Due to the continuous development of new cancer treatments and the sophistication of existing radiotherapy, it has become increasingly challenging to identify the best treatment for a specific patient [1]. A multifactorial clinical decision support systems (CDSS) could help meet this challenge when combining clinical, dosimetric and cost variables (e.g. information about the patient or tumour) with expert knowledge (e.g. on a specific treatment modality) to make a quantitative treatment comparison [2-7]. Such a tool would facilitate individualised radiotherapy treatment.

Given its favourable dose distribution, proton therapy is expected to be less toxic and more effective than photon therapy [8-10]. As a result, many oncology centres around the world have introduced proton therapy over the last decade [11]. However, planning studies show that not all patients would benefit from this more expensive treatment [12,13]. Clinical data-exchange platforms have been developed previously to justify patient stratification for a fair and efficient use of the treatment [14-16]. However, its cost-effectiveness compared to conventional photon radiotherapy is yet unevaluated for many cancers [17-19].

Dutch health authorities have agreed upon the need for a model-based indication methodology to select patients eligible for proton radiotherapy [20-22]. Supplementary Fig. 1 illustrates a Dutch decision tree regarding proton therapy reimbursement. It determines whether a patient is expected to benefit sufficiently from proton therapy justifying reimbursement of the treatment costs. For an effective and efficient evaluation of these aspects, a CDSS is needed that supports the claim whether or not proton therapy is expected to have a clinical benefit in a given patient.

We postulate that such a CDSS should have at least three levels. The first *dosimetric level* should evaluate whether a radiotherapy plan meets predefined dosimetric threshold for a patient's organs at risk (OARs). The second *toxicity level* should estimate whether the probability of radiation induced normal tissue toxicity for the patient is different between different treatment plans. The third *cost-effectiveness level* should evaluate if the extra costs for a certain increase in effectiveness does not exceed a threshold set by society. The effectiveness is defined in quality-adjusted life years (QALYs), which are calculated by estimating the quality and quantity of life extended by a medical intervention [23].

To this end, we developed an online, three-level photon vs. proton CDSS prototype named PRODECIS (PROton DECIsion Support). In this study, we

evaluated the system's performance for patients with head and neck cancer (HNC). Data is provided online on www.cancerdata.org [24].

Materials and Methods

We designed a modular CDSS (Fig. 1) to support the decision between proton and photon therapy. The system was implemented in Java SE 7 (Oracle, Redwood Shores, CA, USA) and Matlab 2010b (Mathworks, Natick, MA, USA) and designed to import photon and proton treatment plans in DICOM-RT format. A PHP webform was created to upload the data and additionally ask for clinical parameters of the complication models. All patient information and results were anonymously stored in a MySQL Workbench 6.0 (Oracle, Redwood Shores, CA, USA) database.

Computation services were separated into three levels. On the dosimetric level, we adopted in-house dose volume histogram (DVH) metrics calculation algorithms to extract mean doses from both photon and proton plans. On the toxicity level, we used a number of validated late toxicity prediction models using the TRIPOD Type 4 standard [25] (e.g., regression models [23,26,27]). On the cost-effectiveness level, we incorporated published Markov models[10] ([23]) to assess the QALY and costs of the treatment.

Experimental setup

To test the system, we used datasets from a ROCOCO cohort of 25 HNC patients for whom both photon and proton plans were available [13]. First, on the dosimetric level we computed the dose to the supraglottis area, the superior pharyngeal constrictor muscle (PCM), and the ipsi- and contralateral parotid glands.

Then, on the toxicity level we estimated the normal tissue complication probability (NTCP) for xerostomia and swallowing dysfunction at 6 and 12 months after therapy, using the models published in previous work [23,26,27]. Since the parotid gland location was indicated with left or right in the given datasets, we defined the ipsi- and contralateral parotid glands as receiving higher or lower doses, respectively.

---

[10] Available on www.predictcancer.org

Finally, on the cost-effectiveness level we used a Markov model constructed for HNC patients [23] with pre-treatment RTOG grade 2- swallowing dysfunction and xerostomia. The model is described in Supplementary Table 1.

Threshold definition

For the purpose of treatment comparison, we collected various thresholds to define clinical benefit. On the dose comparison level, from expert opinions and literature, we defined a clinical benefit when a plan met clinical, desirable OAR mean dose thresholds being parotid gland < 26 Gy, superior PCM < 50 Gy and supraglottis area < 50 Gy [28-30].

On the toxicity level, based on the CTCAEv4.0 toxicity criteria, we considered clinical benefit as a predicted reduction in probability of grade 2+ toxicity of >10%. In addition, we used the definition of a "complication profile" where, for each patient, the toxicity probability reductions exceeding 5% were summed and clinical benefit was set at a total reduction of 15% or more [31].

On the cost-effectiveness level, we set the acceptable cost per additional QALY derived from the Markov model at €80,000. This is the official threshold proposed in the Netherlands by the Dutch Council for Public Health and Care [32].

Statistics

We used two-tailed Wilcoxon signed rank tests to determine whether the differences between plans were significant. P-values of less than 0.05 were considered significant.

Results

System development

The PRODECIS prototype was successfully built on a pipelined image processing framework [33] from within our institute. For scaling purposes, each level of computations was encapsulated into a module and was then installed identically in two parallel pipelines (A and B in Fig. 1). After the whole plan of a treatment was transferred, the respective pipeline began computing. Once both computation pipelines were done, the results were delivered to the third pipeline, comparing the multilevel results with the defined threshold per level. Final-

ly, the comparison results were emailed back to the user. From the 25 datasets, the calculations did not succeed for two, due to DICOM compatibility issues. For every patient, it took approximately five minutes for a computer with standard specifications (Intel® Core™ i5-3210M CPU processor with 2.5GHz, 4 GB memory) to finish all given tasks.
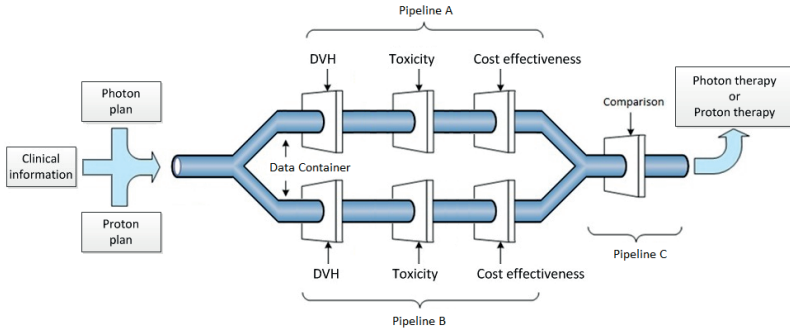
Figure 1 A visualization of the pipeline system which consists of 3 major pipelines.

Experiment results

The system proved successful in the automatic evaluation of proton treatment eligibility according to the model-based approach and predefined thresholds. The number of cases where proton therapy ranked higher as well as average outcomes for both modalities are summarised in Supplementary Table 2.
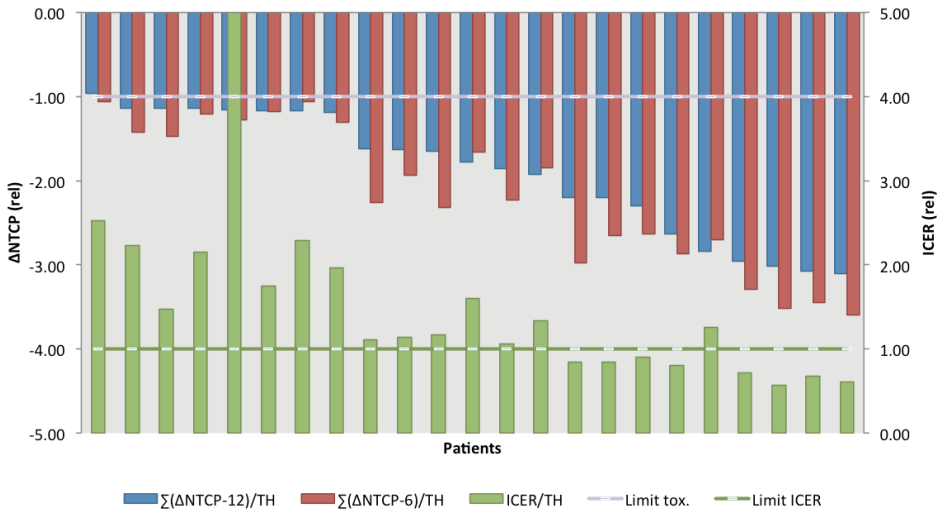
Figure 2 Relative (to threshold of 15%) "complication profiles" after 12 and 6 months (left axis). Only complication predictions that were larger than 5% were included. Right axis for relative (to threshold of €80,000) ICER (=Δ€/ΔQALY).

In Figure 2, the individual results are shown for toxicity and cost-effectiveness, relative to the defined thresholds.

On the dosimetric level, proton therapy significantly lowered doses to the OARs, except for the superior PCM. For the latter, only the proton plans stayed below the thresholds for 2 cases, whereas these were 4, 5 or 12 when considering the supraglottis area, ipsi- or contralateral glands, respectively.

On the toxicity level, proton therapy significantly reduced all toxicities. On average, the probability of swallowing dysfunction 6 months after treatment was reduced from 37% to 28% and from 23% to 18% at 12 months. The probability of xerostomia was reduced for all 23 cases after treatment: from 48% to 25% at 6 months and from 46% to 23% at 12 months. With combined toxicity thresholds, protons outperformed photons for 23 cases at 6 months and 21 cases at 12 months.

On the cost-effectiveness level, we observed an increase in QALY for all the patients in their proton therapy plans, although it was also significantly more expensive. Using the nationally accepted criterion of €80.000 per QALY gained, proton therapy was found to be cost-effective for 8 of the 23 patients.

Discussion

We successfully developed and evaluated the PRODECIS prototype to comply with the Dutch model and added the option to evaluate cost-effectiveness. The study shows that, given nationally accepted guidelines for 15% reduction of a complication profile including swallowing dysfunction and xerostomia, all patients would benefit from proton therapy after 6 months and 91% after 12 months, while 35% would be considered cost-effective at a threshold of 80,000€ per gained QALY. Although a CDSS was previously applied [34,35], we have not found any application that could make quantitative decision-making about photon vs. proton therapy at three levels.

A key characteristic of the system is its parallel pipeline structure, which allows easy extension by reusing the modular code. Another important feature is the dynamic selection of models based on the tumour type. Such flexibility enables the system to rapidly adapt to different user requests and incorporate new insights from the oncology society. Provided the availability of relevant prediction models, future studies could perform systematic experiments to search for an optimal outcome among multiple treatment options at any anatomical site.

The third system feature is its use of the Markov model. It consists of health stages in terms of toxicity RTOG grade and translates toxicity probabilities into transitions between health stages (Supplementary Fig. 2). Through the transition of a patient's health state after treatment, the model estimates the costs and effects of the treatment. An adapted version, referred to as *micro-simulation*, was developed to predict survival of patients with non-small-cell lung cancer [36]. A recently published study [37] shows the same approach to estimate cost-effectiveness of the use of spacers when treating prostate cancer.

Multiple advantages of using the PRODECIS CDSS are foreseen. First, it provides the opportunity for a clinician to make a model-based decision following the Dutch guidelines. Second, it allows clinics to quantitatively prioritise the limited treatment slots and allocate them to the patients expected to gain the most from proton therapy [17,38]. Third, it quantifies clinical evidence for health insurance policy development. Furthermore, it can help in evaluating the cost-effectiveness of deploying a new technology. A final point to note is that consent for data-exchange to the proposed online system can readily be asked from the patient who is being considered for proton therapy and has a direct benefit of the re-use of its data.

However, these advantages will only be achieved when the following concerns are addressed sufficiently. As the system is still in an early stage, extension to a fully operational system offering user management is required to account for audit trails, for instance. As with the MISTIR platform, security measures for encrypted data transfer are to be provided [14]. Furthermore, the system offers a single-shot evaluation and currently lacks proper case management to retrieve previous comparison outcomes for re-evaluation.

Similarly to the ReCompare system [15], the PRODECIS platform is targeted towards referring photon therapy centres, accepting previously calculated photon treatment plans for comparisons. However, PRODECIS also uses user-provided proton plans, whereas ReCompare uses the proton plan generated by the operating proton therapy centre (PTC) itself. Such service can optionally be provided by the staff backing PRODECIS, but as the comparison is performed automatically using standardised models and thresholds an independent evaluation of plan quality and prediction of complication rates is offered to other PTC's as well.

A prerequisite for the automatic numerical evaluation of PRODECIS is that the uploaded treatment plans should adhere to strict protocols, specifying contours per tumour group. The use of so-called "umbrella protocols" and interna-

tional naming convention guidelines will facilitate data exchange in a reusable fashion [39-42]. Although quality assurance methods are implemented, such as contour name mapping, major violation of the protocol will prohibit evaluation, requiring corrections by the user.

It is foreseen that with current efforts from both community as well as industry, instead of calculations based on user-provided treatment plans, automatic plan generation could be applied [43,44]. As an alternative to automatic planning, the estimation of DVH parameters might be reliable and fast, given a sufficiently large historical database of patients with the best planning (as in the study of photon therapy plan optimisation [45]). Patient-selection using a comprehensive matching mechanism based on essential patient characteristics including clinical aspects, tumour location and organ distribution is considered to be incorporated into the PRODECIS system as shown in Supplementary Fig. 3. This will greatly improve the workflow, avoiding the resource-intensive bottleneck of double treatment planning.

A critical factor of the model-based selection method is the quality of the treatment plans under evaluation. Therefore, we expect realistic clinical-grade (thus not "beyond-state-of-the-art") treatment plans that would be administered to the patient in real practice. This means planning protocols need to be up-to-date and in line with the technical possibilities of the treatment options. As for the experiment, the published proton plans for 25 HNC patients are not considered current standard anymore. We have now produced robust treatment plans (unpublished data), which produce not dramatically different but more realistic proton plans, where in some cases the differences are clinically relevant. To further evaluate the system, we will experiment with external datasets from different centres using different treatment techniques.

Furthermore, in the current prototype, the system only considered those toxicities for which reliable NTCP-models were available and that connected to the cost-effectiveness model. Additional models can easily be added including more OARs such as oral cavity, brainstem, or area postrema to predict acute and late radiation-induced toxicities, which may likely be reduced by proton therapy as well and could mean that the cost-effectiveness of protons will be underestimated. For instance, the first comparison of IMRT versus IMPT among oropharyngeal cancer patients treated with chemo-radiation in the MD-Anderson Cancer centre [46] showed a significant decrease of required tube feeding during the course of radiation when IMPT was used. In this regard, direct measurement of QALY's in prospective data registration programs is needed to obtain better

insight into the cost-effectiveness of protons. In addition, to maximise system utility, it is highly desirable to use toxicity models that consider multiple stages. E.g. a reduction of grade 4-5 toxicities is of utmost clinical benefit, but the number of patients is too low to train such a model reliably, which requires international data pooling or rather distributed learning systems [47-49].

The HNC Markov model adopted in this system depended on acceptable costs, which vary from country to country and even from hospital to hospital. It also depended on toxicity estimation models that were regressed without patients' biomedical data. Furthermore, previous interventions such as surgery or chemotherapy were not included in the system, which will bias the complication predictions. Therefore, service at this level is a proof of principle and not conclusive.

The next step will be to include genetic biomarkers of radio-sensitivity to further improve the prediction of late toxicities [50,51]. We aim to continuously update the system with additional models that apply to other diseases and are scalable to other countries. Finally, patient-specific data such as molecular information, patient-reported outcomes and personal preference should be incorporated to truly improve the level of personalisation in decision support systems.

Acknowledgements

References

[1]   Grutters JP, Kessels AG, Pijls-Johannesma M, De Ruysscher D, Joore MA, Lambin P. Comparison of the effectiveness of radiotherapy with photons, protons and carbon-ions for non-small cell lung cancer: a meta-analysis. Radiother Oncol 2010;95:32-40.

[2]   Lambin P, Roelofs E, Reymen B, et al. 'Rapid Learning health care in oncology' - an approach towards decision support systems enabling customised radiotherapy'. Radiother Oncol 2013;109:159-164.

[3]   Pijls-Johannesma M, Grutters JP, Verhaegen F, Lambin P, De Ruysscher D. Do we have enough evidence to implement particle therapy as standard treatment in lung cancer? A systematic literature review. Oncologist 2010;15:93-103.

[4]   Peeters A, Grutters JP, Pijls-Johannesma M, et al. How costly is particle therapy? Cost analysis of external beam radiotherapy with carbon-ions, protons and photons. Radiother Oncol 2010;95:45-53.

[5]   Lambin P, Petit SF, Aerts HJ, et al. The ESTRO Breur Lecture 2009. From population to voxel-based radiotherapy: exploiting intra-tumour and intra-organ heterogeneity for advanced treatment of non-small cell lung cancer. Radiother Oncol 2010;96:145-152.

[6]   Lambin P, van Stiphout RG, Starmans MH, et al. Predicting outcomes in radiation oncology--multifactorial decision support systems. Nat Rev Clin Oncol 2013;10:27-40.

[7]   Lambin P, Zindler J, Vanneste B, et al. Modern clinical research: How rapid learning health care and cohort multiple randomised clinical trials complement traditional evidence based medicine. Acta Oncol 2015:1-12.

[8]   Proton radiotherapy. Horizon scanning report. Health Council of the Netherlands, 2009. Available from: www.gezondheidsraad.nl/sites/default/files/summary_200917.pdf.

[9]   Pijls-Johannesma M, Grutters JP, Lambin P, Ruysscher DD. Particle therapy in lung cancer: where do we stand? Cancer Treat Rev 2008;34:259-267.

[10] Grutters JP, Abrams KR, de Ruysscher D, et al. When to wait for more evidence? Real options analysis in proton therapy. Oncologist 2011;16:1752-1761.

[11] Jermann M. Particle Therapy Statistics in 2014. Int J Part Ther 2015;2:50-54.

[12] Roelofs E, Engelsman M, Rasch C, et al. Results of a multicentric in silico clinical trial (ROCOCO): comparing radiotherapy with photons and protons for non-small cell lung cancer. J Thorac Oncol 2012;7:165-176.

[13] van der Laan HP, van de Water TA, van Herpt HE, et al. The potential of intensity-modulated proton radiotherapy to reduce swallowing dysfunction in the treatment of head and neck cancer: A planning comparative study. Acta Oncol 2013;52:561-569.

[14] Roelofs E, Persoon L, Qamhiyeh S, et al. Design of and technical challenges involved in a framework for multicentric radiotherapy treatment planning studies. Radiother Oncol 2010;97:567-571.

[15] Luhr A, Lock S, Roth K, et al. Concept for individualized patient allocation: ReCompare--remote comparison of particle and photon treatment plans. Radiat Oncol 2014;9:59.

[16] Lock S, Roth K, Skripcak T, et al. Implementation of a software for REmote COMparison of PARticlE and photon treatment plans: ReCompare. Z Med Phys 2015;25:287-294.

[17] Bekelman JE, Asch DA, Tochner Z, et al. Principles and reality of proton therapy treatment allocation. Int J Radiat Oncol Biol Phys 2014;89:499-508.

[18] Pijls-Johannesma MC, de Ruysscher DK, Dekker AL, Lambin P. [Protons and ions in the treatment of cancer; a systematic review of the literature]. Ned Tijdschr Geneeskd 2006;150:2435-2441.

[19] Grutters JP, Pijls-Johannesma M, Ruysscher DD, et al. The cost-effectiveness of particle therapy in non-small cell lung cancer: exploring decision uncertainty and areas for future research. Cancer Treat Rev 2010;36:468-476.

[20] Langendijk JA, Lambin P, De Ruysscher D, Widder J, Bos M, Verheij M. Selection of patients for radiotherapy with protons aiming at reduction of side effects: the model-based approach. Radiother Oncol 2013;107:267-273.

[21] Widder J, van der Schaaf A, Lambin P, et al. The quest for evidence for proton therapy: the model-based approach and precision medicine. Int J Radiat Oncol Biol Phys 2015.

[22] Grau C. The model-based approach to clinical studies in particle radiotherapy--a new concept in evidence based radiation oncology? Radiother Oncol 2013;107:265-266.

[23] Ramaekers BL, Grutters JP, Pijls-Johannesma M, Lambin P, Joore MA, Langendijk JA. Protons in head-and-neck cancer: bridging the gap of evidence. Int J Radiat Oncol Biol Phys 2013;85:1282-1288.

[24] Cheng Q, Roelofs E, Ramaekers B, et al. Data from: Development and Evaluation of an Online Three-Level Proton vs Photon Decision Support Prototype for Head and Neck Cancer - Comparison of Dose, Toxicity and Cost-Effectiveness. CancerData. 2015. doi:10.17195/candat.2015.10.5.

[25] Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. Eur J Clin Invest 2015;45:204-214.

[26] Christianen ME, Schilstra C, Beetz I, et al. Predictive modelling for swallowing dysfunction after primary (chemo)radiation: results of a prospective observational study. Radiother Oncol 2012;105:107-114.

[27] Beetz I, Schilstra C, Burlage FR, et al. Development of NTCP models for head and neck cancer patients treated with three-dimensional conformal radiotherapy for xerostomia and sticky saliva: the role of dosimetric and clinical factors. Radiother Oncol 2012;105:86-93.

[28] Deasy JO, Moiseenko V, Marks L, Chao KS, Nam J, Eisbruch A. Radiotherapy dose-volume effects on salivary gland function. Int J Radiat Oncol Biol Phys 2010;76:S58-63.

[29] Rancati T, Schwarz M, Allen AM, et al. Radiation dose-volume effects in the larynx and pharynx. Int J Radiat Oncol Biol Phys 2010;76:S64-69.

[30] Levendag PC, Teguh DN, Voet P, et al. Dysphagia disorders in patients with cancer of the oropharynx are significantly affected by the radiation therapy dose to the superior and middle constrictor muscle: a dose-effect relationship. Radiother Oncol 2007;85:64-73.

[31] Consensus document voor selectie van patiënten met een model-based indicatie voor protonen therapie (Unpublished internal document). Landelijk Platform Protonen Therapie (LPPT), 2015.

[32] Sensible and sustainable care. Council for Public Health and Care (RVZ), 2006. Available from: www.raadrvs.nl/uploads/docs/Sensible_and_sustainable_care.pdf.

[33] Van Soest J, Lustberg T, Grittner D, et al. Towards a semantic PACS: Using Semantic Web technology to represent imaging data. Stud Health Technol Inform 2014;205:166-170.

[34] Seroussi B, Soulet A, Messai N, Laouenan C, Mentre F, Bouaud J. Patient clinical profiles associated with physician non-compliance despite the use of a guideline-based decision support system: a case study with OncoDoc2 using data mining techniques. AMIA Annu Symp Proc 2012;2012:828-837.

[35] Steele SR, Bilchik A, Johnson EK, et al. Time-dependent estimates of recurrence and survival in colon cancer: clinical decision support system tool development for adjuvant therapy and oncological outcome assessment. Am Surg 2014;80:441-453.

[36] Bongers ML, De Ruysscher D, Oberije C, Lambin P, Uyl-de Groot CA, Coupe VM. Calibration And Statistical Modeling To Inform A Micro-Simulation Model For Early HTA. *Value in Health* 2014;17(7):A561.

[37] Vanneste BG, Pijls-Johannesma M, Van De Voorde L, et al. Spacers in radiotherapy treatment of prostate cancer: Is reduction of toxicity cost-effective? Radiother Oncol 2015;114:276-281.

[38] Jakobi A, Bandurska-Luque A, Stutzer K, et al. Identification of Patient Benefit From Proton Therapy for Advanced Head and Neck Cancer Patients Based on Individual and Subgroup Normal Tissue Complication Probability Analysis. Int J Radiat Oncol Biol Phys 2015;92:1165-1174.

[39] CancerData.org. EuroCAT Umbrella Protocol for NSCLC 2015 [cited 2015 6 Oct]. Available from: www.cancerdata.org/protocols/eurocat-umbrella-protocol-nsclc.

[40] Meldolesi E, van Soest J, Dinapoli N, et al. An umbrella protocol for standardized data collection (SDC) in rectal cancer: a prospective uniform naming and procedure convention to support personalized medicine. Radiother Oncol 2014;112:59-62.

[41] Melidis C, Bosch WR, Izewska J, et al. Global harmonization of quality assurance naming conventions in radiation therapy clinical trials. Int J Radiat Oncol Biol Phys 2014;90:1242-1249.

[42] Santanam L, Hurkmans C, Mutic S, et al. Standardizing naming conventions in radiation oncology. Int J Radiat Oncol Biol Phys 2012;83:1344-1349.

[43] Voet PW, Dirkx ML, Breedveld S, Fransen D, Levendag PC, Heijmen BJ. Toward fully automated multicriterial plan generation: a prospective clinical study. Int J Radiat Oncol Biol Phys 2013;85:866-872.

[44] Kraan AC, van de Water S, Teguh DN, et al. Dose uncertainties in IMPT for oropharyngeal cancer in the presence of anatomical, range, and setup errors. Int J Radiat Oncol Biol Phys 2013;87:888-896.

[45] Petit SF, Wu B, Kazhdan M, et al. Increased organ sparing using shape-based treatment plan optimization for intensity modulated radiation therapy of pancreatic adenocarcinoma. Radiother Oncol 2012;102:38-44.

[46] Frank SJ, Rosenthal DI, Ang K, et al. Gastrostomy Tubes Decrease by Over 50% With Intensity Modulated Proton Therapy (IMPT) During the Treatment of Oropharyngeal Cancer Patients: A Case–Control Study. Proceedings of the American Society for Radiation Oncology 55th Annual Meeting ASTRO's 55th Annual Metting 2013;87.

[47] Deasy JO, Bentzen SM, Jackson A, et al. Improving normal tissue complication probability models: the need to adopt a "data-pooling" culture. Int J Radiat Oncol Biol Phys 2010;76:S151-154.

[48] Dekker A, Nalbantov G, Oberije C, et al. Multi-centric learning with a federated IT infrastructure: application to 2-year lung-cancer survival prediction. Radiother Oncol 2013 Mar;106:S193.

[49] Skripcak T, Belka C, Bosch W, et al. Creating a data exchange strategy for radiotherapy research: towards federated databases and anonymised public datasets. Radiother Oncol 2014;113:303-309.

[50] Fachal L, Gomez-Caamano A, Barnett GC, et al. A three-stage genome-wide association study identifies a susceptibility locus for late radiotherapy toxicity at 2q24.1. Nat Genet 2014;46:891-894.

[51] Barnett GC, Thompson D, Fachal L, et al. A genome wide association study (GWAS) providing evidence of an association between common genetic variants and late radiotherapy toxicity. Radiother Oncol 2014;111:178-185.
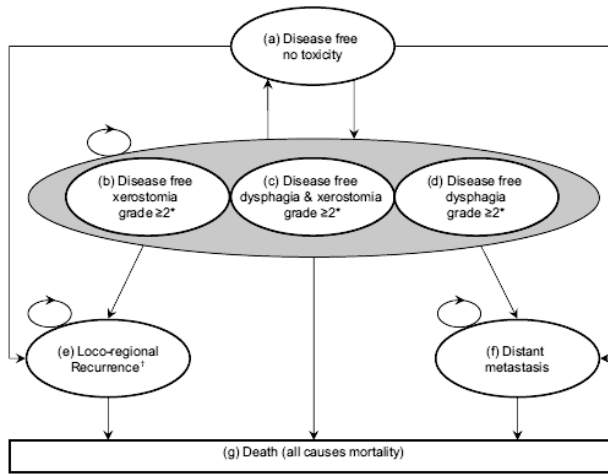
Supplementary Material

Table A1 Safety margin as a function of the setup error and the peak-to-peak breathing amplitude (in mm).

| Setup error	Breathing amplitude	Safety margin
5	0	5
5	5	6
5	10	7
5	20	8
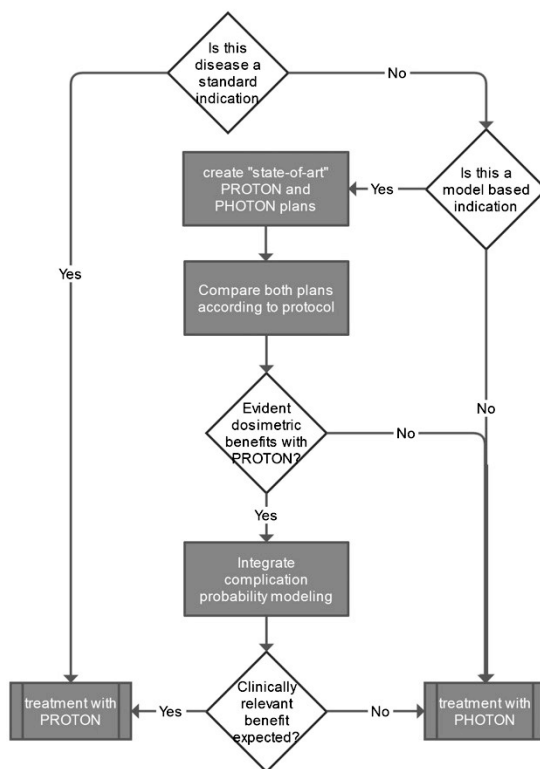
Supplementary Figure 1 The decision tree for all model-based indications to get proton therapy reimbursement.

Supplementary Figure 2 A cost effectiveness model in terms of Markov Cycle Tree [1]. "Toxicity that occurred in the first 6 months was (partly) reversible. Therefore, patients are allowed to move between health states a, b, c, and d 6 months after radiation therapy. Thereafter, toxicity was assumed to be irreversible.

*Toxicity was defined according to the presence of toxicities grade 2 or higher (RTOG)."

Reprinted from Green Journal, Vol. 85, Ramaekers BL et al., Protons in head-and-neck cancer: bridging the gap of evidence, pp. 1282, Copyright (2013), with permission from Elsevier.
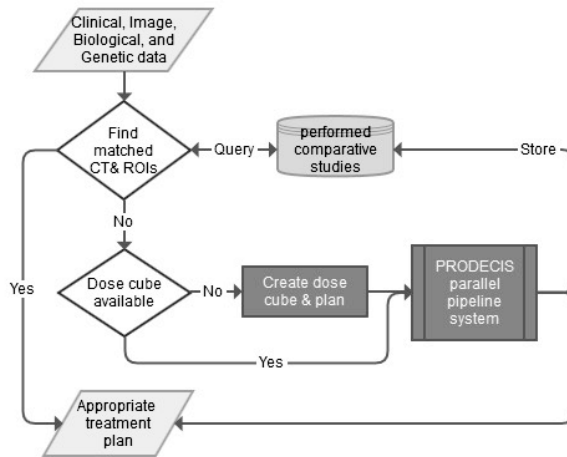
Supplementary Figure 3 A future perspective: incorporating matching mechanism into the PRODECIS system.

Supplementary Table 1 NTCP models and Cost-effectiveness estimation model [1]

NTCP models *	Coefficient	Se
Probability RTOG ≥grade 2 xerostomia 6 months after radiotherapy (N=354)		
Constant	-2.852	0.313
Mean dose to parotis ipsilateral	0.031	0.009
Mean dose to parotis contralateral	0.036	0.010
Probability RTOG ≥grade 2 xerostomia 12 months after radiotherapy (N=251)		
Constant	-2.949	0.386
Mean dose to parotis ipsilateral	0.025	0.011
Mean dose to parotis contralateral	0.045	0.012
Probability RTOG ≥grade 2 dysphagia 6 months after radiotherapy (N=354)		
Constant	-6.094	0.925
Mean dose to pharyngeal constrictor muscle superior	0.057	0.009
Mean dose to supraglottic area	0.037	0.012
Probability RTOG ≥grade 2 dysphagia 12 months after radiotherapy (N=230)		
Constant	-6.343	1.186
Mean dose to pharyngeal constrictor muscle superior	0.058	0.012
Mean dose to supraglottic area	0.027	0.015
Cost-effectiveness estimation	**Coefficient**	
Cost estimation for IMRT		
Constant	35827€	
6 months after radiotherapy		
RTOG ≥grade 2 xerostomia and RTOG ≥grade 2 dysphagia	1777€	
RTOG ≥grade 2 xerostomia without RTOG ≥grade 2 dysphagia	24€	
RTOG ≥grade 2 dysphagia without RTOG ≥grade 2 xerostomia	1753€	
12 months after radiotherapy		
RTOG ≥grade 2 xerostomia and RTOG ≥grade 2 dysphagia	19418€	
RTOG ≥grade 2 xerostomia without RTOG ≥grade 2 dysphagia	299€	
RTOG ≥grade 2 dysphagia without RTOG ≥grade 2 xerostomia	19119€	
Cost estimation for IMPT		
Constant	46880€	
6 months after radiotherapy		
RTOG ≥grade 2 xerostomia and RTOG ≥grade 2 dysphagia	1777€	
RTOG ≥grade 2 xerostomia without RTOG ≥grade 2 dysphagia	24€	
RTOG ≥grade 2 dysphagia without RTOG ≥grade 2 xerostomia	1753€	
12 months after radiotherapy		
RTOG ≥grade 2 xerostomia and RTOG ≥grade 2 dysphagia	19418€	
RTOG ≥grade 2 xerostomia without RTOG ≥grade 2 dysphagia	299€	
RTOG ≥grade 2 dysphagia without RTOG ≥grade 2 xerostomia	19119€	
QALY estimation for both	**Coefficient**	
Constant	6.79	
6 months after radiotherapy		
RTOG ≥grade 2 xerostomia and RTOG ≥grade 2 dysphagia	-0.049	
RTOG ≥grade 2 xerostomia without RTOG ≥grade 2 dysphagia	-0.022	
RTOG ≥grade 2 dysphagia without RTOG ≥grade 2 xerostomia	-0.032	
12 months after radiotherapy		
RTOG ≥grade 2 xerostomia and RTOG ≥grade 2 dysphagia	-0.767	
RTOG ≥grade 2 xerostomia without RTOG ≥grade 2 dysphagia	-0.346	
RTOG ≥grade 2 dysphagia without RTOG ≥grade 2 xerostomia	-0.497	

Abbreviations: Se = standard error, NTCP = normal tissue complication probability, IMRT = intensity-modulated radiotherapy with photons, IMPT = intensity-modulated proton radiotherapy.

* The multivariate normal distribution was constructed using Cholesky decompositions [2]

Supplementary Table 2 Comparison statistics on 23 HNC patients, showing the number of patients that were expected to benefit from proton therapy, based on the given thresholds. Furthermore, the average OAR mean doses and significances for differences between the groups are shown.

	Photon plans		Proton plans		H_0	Compliance*
Dose statistics						
Endpoints (Gy)	Mean	SD	Mean	SD	P-value	Cases (TH)
Mean dose to supraglottis area	54.0	12.1	39.9	25.6	<0.001	4 (<50 Gy)
Mean dose to superior PCM	62.2	3.6	61.1	5.1	0.465	2 (<50 Gy)
Mean dose to ipsi-lat. parotid gland	47.1	11.0	30.1	9.5	<0.0001	5 (<26 Gy)
Mean dose to contra-lat. parotid gland	35.8	9.6	20.9	6.3	<0.0001	12 (<26 Gy)
Toxicity statistics						
Endpoints (%)	Mean	SD	Mean	SD	P-value	Cases (TH**)
6 months after treatment						
Swallowing dysfunction	37.3	12.1	28.4	17.1	<0.001	13 (Δ>10%)
Xerostomia	47.7	13.4	24.7	8.1	<0.0001	22 (Δ>10%)
Combined †						23 (Δ>15%)
12 months after treatment						
Swallowing dysfunction	22.6	7.8	17.9	10.0	<0.001	1 (Δ>10%)
Xerostomia	46.3	13.9	23.1	7.7	<0.0001	22 (Δ>10%)
Combined †						21 (Δ>15%)
Cost-effectiveness statistics						
Endpoints	Mean	SD	Mean	SD	P-value	Cases (TH)
Cost (€)	40970	1692	50874	2208	<0.0001	
QALY	6.51	0.06	6.62	0.06	<0.0001	
ICER (Δ€/ΔQALY) ‡	Mean: 118546		SD: 76605			8 (<80000)

Abbreviations: PCM: pharyngeal constrictor muscle, QALY: Quality-Adjusted Life Year, SD: standard deviation, TH: thresholds.

* the number of patient cases where only the proton plan was favourable given the threshold;
** absolute reduction (photon – proton) in toxicity prediction (%)
† only toxicity reductions exceeding 5% were summed up and should exceed 15%
‡ Incremental Cost-Effectiveness Ratio: incremental costs divided by incremental QALYs.

References

1. Ramaekers BL, Grutters JP, Pijls-Johannesma M, Lambin P, Joore MA, and Langendijk JA Protons in head-and-neck cancer: bridging the gap of evidence. Int J Radiat Oncol Biol Phys 2013; 85(5): 1282-8.
2. Briggs A SM, Claxton K Decision Modelling for Health Economic Evaluation. Oxford: Oxford University Press 2006.
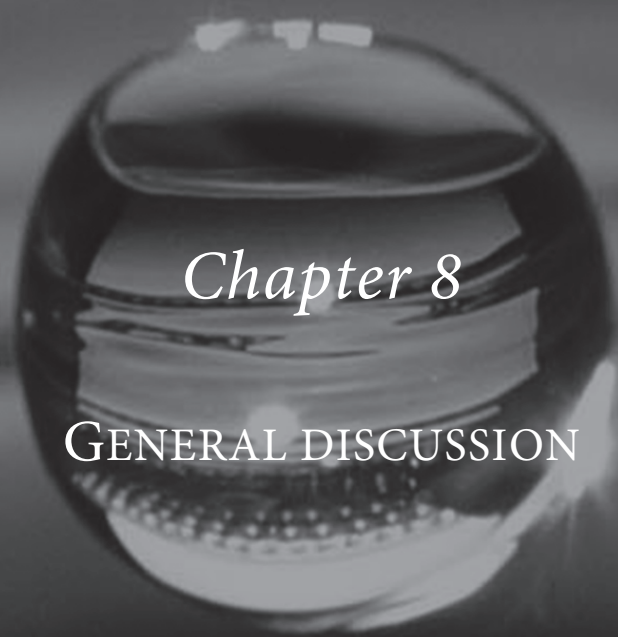
Chapter 8

GENERAL DISCUSSION

Introduction

This thesis shows the merger of two seemingly distinct topics: knowledge engineering and proton therapy. However, in retrospect, the research shows in a logical transition that the application of proton therapy greatly depends on worldwide knowledge previously gained and how new methods to transparently mine, learn and distribute that knowledge contribute to the realisation of cost-effective healthcare even when expensive technology such as proton therapy is considered.

The central theme in this work is to investigate the differences between patients and to learn how to choose the best treatment on an individual basis, with the assistance of multi-modal prediction models in clinical decision support systems.

The specific aims for this thesis were 1) to attest the increased need for information exchange and knowledge extraction on an international level to aid clinical decision-making for improved individualized healthcare, 2) investigating supplemental evidence for the justification of giving more expensive proton therapy to those patients that are expected to benefit the most by means of *in silico* trials and finally 3) to show the realisation of the merger of the two research areas with the implementation of a reimbursement eligibility decision aid for proton therapy.

Challenges in oncology

Controlling patient outcome

One of the many challenges in oncology is to handle the great diversity in patients with different characteristics such as age, clinical history or concomitant treatments. On top of that, there is a huge heterogeneity in tumour types, depending on genetic and micro-environmental circumstances, for instance, which influences the reaction to treatment [1]. Furthermore, not only the information that is available to classify patients is exploding due to advancing techniques in imaging and genomic sequencing, for instance, but also the options for treatment are ever increasing. This leaves the acting physician with a huge challenge to combine all this information and find the best treatment option.

The discipline of systems and control theory teaches us that the best way to steer a complex system towards a desired output is by implementing an active feedback loop using direct measurement from the endpoint that needs to be modified. However, in the treatment process, there are also many intermediate

'endpoints' that need to be controlled such as a patient's position and given dose. As shown in Figure 7, these are acting as cascaded loops that differ in speed of interaction. The inner loop is fast with for instance daily imaging before the treatment is given [2]. The outer loop uses a slower feedback, measuring dose during treatment [3–6]. A change in the treatment is usually only after a few fractions unless major disturbances are experienced.
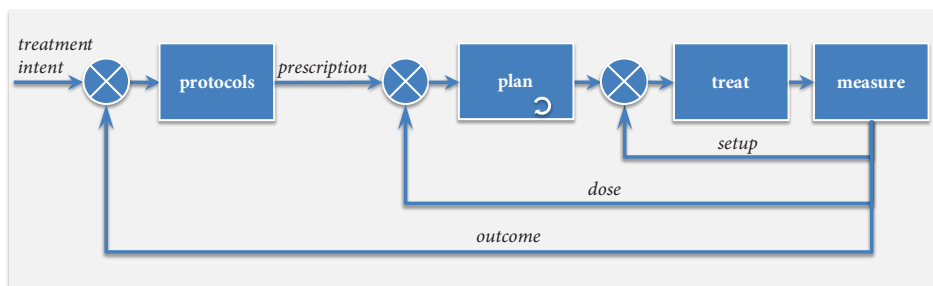
Figure 7 Cascade control of individual treatment quality

Analogous to this, to improve the outcome of cancer patients, we must measure and use patient outcome in an encompassing cascaded loop. However, collection of outcome data is a burden as it lacks standardisation and unfortunately is hampered by the increasingly competitive environment in which healthcare is operating nowadays [7–10].

A recently published report by the National Audit Office showed the failure of the UK Department of Health and National Health Service (NHS) England to collect outcome data of cancer patients receiving chemotherapy through the Cancer Drugs Fund (CDF) [11,12]. The Fund realised the importance of collecting outcome and toxicity data from the start but after having spent nearly 1 billion pounds during 4.5 years, only an astonishing low number (7%) of the records had outcome summaries.

Classical learning cycle

The standard method to improve cancer care is through the generation of "Level I evidence" via randomised controlled trials. It is known, however, that generally the number of cancer patients participating in trials is less than 5%, with some exceptions in specific trial programs [13–16]. Furthermore, the patients eligible for participation in trials are carefully selected for age, health performance score and previous disease. This decreases the chance considerably that a patient in

need of an optimal treatment fits the profile of the trial population that led to the discovery of this treatment. How can we then be certain about the efficacy of this treatment?

Furthermore, with late side effects or clinical outcome as main endpoint, the turnaround time to generate new protocols from the results of clinical trials is in the order of five to ten years, or even more [17]. Given the rapid development of new treatment equipment and continuous biological discoveries the classical approach of evidence-based medicine should be considered insufficient to keep up.

This is for instance the case with proton therapy, which is a cancer treatment modality that is in a very active development stage, currently. For instance, real-time imaging of patient position and measurement of actual delivered dose, which is common in the photon therapy field, is now rapidly being introduced in the proton world. In past years, many technical improvements have been introduced and not all of them underwent a rigorous approval through clinical trials. One can think of the use of blocks, wedges and multi-leaf collimators instead of square fields. However, when the expected gain is unclear and costs are considerable, there is a strong need of an alternative method for proof finding [18,19].

Exploring frontiers with Rapid Learning

The general concept of Rapid Learning is to complement the evidence-based medicine approach for clinical decision-making with prediction modelling and decision support tooling in an iterative loop (Chapter 2, [20–22]). One could think of this approach as being a Deming "**Plan-Do-Study-Act**" (PDSA[11]) cycle where standing clinical practice is continuously updated with new insights leading to improved treatment processes with better outcomes.

Deming introduced the PDSA cycle, also known as the Deming wheel, which offers a method for continuous improvement through small, iterative steps [23,24]. Originally targeted towards optimising quality in the manufacturing industry, the PDSA method can equally well be applied to other fields, such as improving quality and safety in healthcare [25,26].

Referring to the CDF program described above, it shows that the efforts undertaken to improve access to cancer drugs are undermined by insufficient *planning* for data collection and *doing* documentation of data and observations.

---

[11] The commonly known PDCA cycle (with *check* replacing *study*) is a Japanese interpretation of the principles taught by Deming in 1950, who considered it a 'corruption' of the original concept.

This leads to flaws in *studying* the results and subsequently limits options for *action* and subsequent improvement cycles.

When treating patients, there are many variables that need to be taken into account, ranging from the time needed for treatment plan preparations, dose to the tumour and normal tissues, to overall treatment time, number of staff involved and the costs of treatment, to name a few. From multi-objective optimisation, e.g. in treatment planning systems, we know that multiple combinations of these variables in this multi-dimensional search space exist that are considered optimal: the *Pareto Frontier* [27–29]. On this frontier, there is no other combination of factors possible that result in a better treatment, is more cost-effective, offers better patient satisfaction, etc. Given all available options and the limitations we are bound by, we need to strive for treatments that satisfy these conditions. Every new option that is added to the realm of cancer treatment techniques will add another dimension to the problem space (or *solution space* for the optimists) and thus needs exploring new frontiers.

To this end, Rapid Learning uses routinely collected clinical data, including historical populations, to predict clinical outcome of new patients and thus facilitate individualised decision-support. We could state that Rapid Learning is a way to empirically measure the coordinates of solutions in the multi-dimensional cancer treatment space and the objective is to find the Pareto effective frontier for individual patients. When we realise that every patient adds unique physical and biological parameters and personal preferences, for instance, it is clear that the problem space drastically expands (or rather explodes) and that it is impossible for anyone (even physicists) to find this solution him- or herself.

This is where Rapid Learning includes the use of models that, given a set of input parameters, predict the outcomes of possible treatments. As indicated in Chapter 1, the models can be dose- and data-driven. In either way, much data is required to generate and validate these models, which underpins the importance of both registration of meaningful information before and along the treatment process, as well as capturing the crucial follow-up data.

Hidden information

The CDF example shows how important it is to have intermediate and end-point measurements when you want to "close the loop" during an intervention. However, in oncology, it is not always easy to find this information. Although it

is common practice to use electronic medical records (EMR) in oncology, its use is primarily 'operations' oriented. The main function that vendors support is the management and documentation of the treatment process itself. By design, EMRs are not well suited for learning systems. They are mostly internally oriented and reuse of the information is difficult, if not impossible without in-house software development.

In Chapter 3, we showed how the information from our own medical record (EMD) needed to be combined with additional information sources and software components for extraction of missing or incorrect information. Although it took considerable time for complete implementation, it was relatively easy to realise because the EMD was an in-house development, which gave full access to the internal data formats and structure. Current EMRs do tend to improve in offering data reports. However, they are often limited in the freedom to collect, combine and report information outside the regular process patterns.

For example, documentation from essential patient reported outcome measures (PROMs) is often not implemented in a standard, qualified manner [7–9,30]. Also, the information that was provided is typically unavailable in the regular interface and therefore the use of PROMs disturbs the workflow. Implementation of electronic patient portals that offer flexibility and self-support by the institution is crucial to make the use of PROMs as learning material possible. Furthermore, there is an increased need to export treatment information to national registries. This 'outer cascade' type of information will help to set and improve field standards for generic treatment quality, such as waiting times or treatment costs.

Frequent reporting of these data still proves to be a burden and direct electronic exchange is only possible after considerable investment in the development of special software. Part of the problem is that certain information has to be provided, while the data are not available in the institution itself. This is typically true in oncology as there are many referring partners where patients originally are diagnosed and radiotherapy can be combined with surgery or chemotherapy. Due to the lack of a national electronic patient record, the provisioning of relevant information from the chain of medical interventions is therefore extremely difficult.

Opportunities in radiotherapy

Data-driven healthcare

With the high number of variables in the treatment of cancer and the need to be more precise while maintaining efficiency we need to be able to accurately predict treatment outcome and use that in daily clinical routine. Luckily, radiotherapy is among the most data-driven disciplines in healthcare using many types of (digital) models to prepare the best personal treatment for its patients. Patients are digitised via CT, MR or PET imaging after which computer models of linear accelerators are used to balance the amount of the dose given to tumour and healthy tissue to strive for a possible cure without complications. During the treatment, repetitive images are captured to correct for misalignment of the patient and detect geometrical changes. Additionally, a technique was developed at MAASTRO clinic that measures the given dose to the patient and prevents over- or under dosage [3–6]. This offers the opportunity to study and minimise the differences between the intended (planned) and given (real) dose.

Given the early adoption of EMRs and many treatment measurements to monitor outcome and control quality there is a wealth of clinical data to extract information from and use for data mining, providing outcome prediction models, for instance. However, as discussed above, this information is often incomplete.

As shown in Chapter 3, it requires substantial effort to bring together the multiple sources of information in a data warehouse and provide methods to automatically generate missing information from the collected data. However, it greatly stimulates structuring the data for reuse and improves quality as well as quantity, which show its importance in the benefit of data collection in clinical trials.

Breaking down barriers with CAT

As previously indicated, data-driven prediction modelling demands a large amount of data and as the patient population is very heterogeneous, the required data should be too. Furthermore, a model should not be validated using the same data as was used for the modelling itself. This has given rise to the need of extramural data collections.

When exchanging patient-related data, there are many barriers to break down. Unfortunately,

> **Barriers to share**
>
> … the problem is not really *technical* … Rather, the problems are *ethical*, *political*, and *administrative*
>
> *Lancet Oncol, 2011*

overly cautious interpretation of European Directives on Good Clinical Practice [31] and Personal Data [32] has had counterproductive effects in increasing cost and subsequent reduction of clinical trials and instalment of national cancer registries and the information exchange of this information [33–35].

In 2005, we started the Computer Aided Theragnostic[12] (CAT) research project (Chapter 3, 4) to address the clinical information exchange issue and facilitate data collection and knowledge transfer by building an international data-sharing infrastructure. Through daily synchronisation of the various clinical databases the data is aggregated into a de-identified, disease-oriented research environment. Via a web-based research portal, data collections can be retrieved for analysis in a common data model. In Chapter 4, we showed a prototype for a special implementation of a CAT system for international collaboration. This system was built using free and open-source software and enabled data exchange of de-identified medical record data and accompanying image datasets.

The EuroCAT project (eurocat.info), which ran from 2010 to 2015 among five institutions in Belgium, Germany and the Netherlands, exploited the CAT principle and built an international, multi-lingual data-sharing network. The project proved successful in automatic data-extraction of predictive parameters for complication prediction models, among others. The EuroCAT project has laid the groundwork for a global infrastructure and subsequent CAT projects that leverage local medical databases with information collected from daily clinical routine and share that data in a secure way.

In time, it will be impossible to continue to collect trial or modelling data in a centralised manner as the digital information that belongs to a patient treatment expands drastically with new 'omics' datasets, for instance. Instead, this requires centres to install semantic interoperable connectors to clinical databases, which enables a global *federated learning network* such that the data does not leave the hospital but lightweight and non-identifying learning applications do have access to those collections. The challenge here is to have the learning models converge, while operating only on data within the silos behind hospital walls. By continuous exchange of machine-learning regression factors to a central server, a consensus model can be derived without violating the patient's privacy [36,37].

---

[12] "Theragnostics" is an aggregation of *therapeutics* and *diagnostics*

Model-based decisions

In 2006, MAASTRO CLINIC was one of the initiators of the first proposal to offer proton therapy to patients in the Netherlands. Although the benefit of the physical advances of using a proton beam when compared to a photon beam is acknowledged for some indications, such as paediatrics, intraocular, chordoma and chondrosarcoma, evidence hereof is still considered insufficient for many other cancer types. To investigate the possible benefit for subgroups of patients, we initiated the Radiation Oncology Collaborative Comparison (ROCOCO) group in 2007. This international research group consists of multiple investigators renowned for their expertise in disease-specific treatment options, radiobiology, physics, treatment planning, etc. in the photon and particle therapy field. We set up a *virtual enterprise* type of trial infrastructure, MISTIR, as described in Chapter 5, which allows secured data transfer and collaboration between multiple international research groups.

We applied the principle of *in silico* simulations of several clinical trials to compare multiple treatment modalities for different cancer patient groups[13]. This method offers the benefit that the same patient is 'treated' several times and acts as its own control in the comparison, which reduces the necessary number of cases for the trial. We investigated differences between various treatment modalities and patient groups, for example between three-dimensional conventional and intensity-modulated photon therapy and passive-scatter proton therapy for long cancer, as shown in Chapter 6. We found that some patient subgroups are likely to benefit from particle therapy but others not.

The principle of performing simulations to test the introduction of a new technology fits the Rapid Learning PDSA approach described earlier. First, the new treatment technique is simulated and multi-factorial complication models predict the expected outcome. Based on national consensus [38], a selection criterion is defined whether a patient is expected to sufficiently benefit from the new technique (e.g. proton therapy). If so, this modality is used to treat the patient, while the related information is entered into a national database to enable prospective cohort studies. Such a Rapid Learning or *model-based* approach is unique in the field of radiotherapy and offers the opportunity to generate clinical evidence of new treatments with rapidly changing techniques [39–41].

---

[13] See www.mistir.info/activities for an overview of past and current trials

Based on this principle, we generated a unique model-based proton therapy decision support system (PRODECIS), as shown in Chapter 7. This online decision aid compares photon and proton therapy using components of the *MISTIR* framework and uses validated prediction models from *PredictCancer* (predictcancer.org) to estimate the complication with both modalities. Additionally, we implemented a model that indicates whether the more expensive proton therapy is cost-effective, taking a nationally accepted price per life year gained with good quality into account [42]. Although initially the model-based approach is implemented for proton therapy in the Netherlands, it is applicable to other treatment modalities as well [43]. It is expected that this method will be used more frequently for healthcare environments with rapidly advancing technical solutions.

Conclusions and future directions

Standardising medical data exchange

Knowledge is power. Unfortunately, we sometimes lack the power to build that knowledge. In a striking report to the US Congress, called "Report on Health Information Blocking", the Office of the National Coordinator for Health Information Technology (ONC) concluded that there is evidence of active blocking of medical information exchange by EMR vendors [44]. This conclusion is harsh and probably not only to be held against the vendors. As mentioned, disturbing interpretation of European Directives has caused considerable barriers to share crucial clinical information as well.

Fortunately, as we have shown, by combining Medical Informatics and Knowledge Engineering we are able to offer a powerful and proven methods to break these barriers down and enable exchange of medical data through active international collaboration to individualise healthcare. This form of *translational informatics* is highly needed, as there is still a huge demand for tooling and regulations to facilitate global data-exchange. Although the discipline of data mining needs a heterogeneous data pool, it cannot compensate all of the data quality issues simply by outnumbering. To this end, harmonisa-

> **" Barriers to share**
>
> … this report suggests that some persons and entities are interfering with the exchange or use of electronic health information …
>
> *ONC, April 2015*

tion of data-standards is required.

We therefore actively support the concept of *Umbrella protocols* with the publication of open-access protocols on CancerData [45–47]. The primary objective of these protocols is to standardise data collection [48]. Other standards to specifically equalise naming conventions in radiation oncology terms have been proposed as well [49,50]. However, to date, no treatment planning system adheres to these conventions.

Although international standards are available, such as ICD, LOINC, NCI Thesaurus and SNOMED Clinical Terms [51–54], their incorporation into EMRs is still limited [55]. We have shown pioneering work of our Clinical Data Science research group in which medical data was extracted from two EMRs and converted the terminology from Dutch and Italian into common SNOMED CT terms (Chapter 4) to enable distributed learning for individualised medicine[14]. The available declarations linking specific common terminology, concepts and relations for radiotherapy are often limited, in these global standards.

Semantic interoperability in ongology

Therefore, a dedicated Radiation Oncology Ontology (ROO) has been developed, which is an ontology to publish Linked Data for radiotherapy. The ROO reusing as much entities as possible from existing ontologies, but adds what is lacking for radiation oncology (cancerdata.org/roo-information). The ROO is published online as an open-access resource to stimulate reuse and standardisation in the field.

By translation of patient data into globally unique concepts, properties and relations, it is possible to create a fully de-identified semantic interoperable representation of this information. This means that computer programs can independently interpret this data and process it for data mining purposes. One project in which we participate as clinical partner that uses such technology is the Semantic DICOM (SeDI) project (semantic-dicom.org), which develops a web-based, real-time queryable image archive [56,57]. It uses a Resource Description Framework (RDF) model to store entity-relationships of image data in *triples* of subject–predicate–object expressions.

This enables the revolutionary concept of Linked (Open) Data (lod2.eu), which will change the Internet as we currently know it from collections of publications into a web of knowledge. As an example, DBpedia (dbpedia.org) uses

---

[14] You can find a video showing the principles here: youtu.be/ZDJFOxpwqEA

this concept to enable a faceted search for expressions such as "Give me all Italian musicians from the 18th century". Such an approach translates into an oncology research domain as a question "Give me 100 female patients with a prostate tumour" would actually be answered with no results, as the system *knows* that a female person does not have a prostate, and not return incorrect information based on the keywords. *Knowledge is power…*

For instance, after playing chess for years, IBM is now putting supercomputing to good use with the Watson project. Using massive computer power, natural language processing, linked data, 'deep learning' techniques and the clinical expertise of several cancer institutes, the Watson Health project is building cognitive artificial intelligence solutions for clinical use against cancer [58–61].

Stimulating reuse with open-access data

The success of rapid learning will greatly depend on the quantity and quality, and public availability of published healthcare data. Semantic web technology can offer solutions to bring high quantities of EMR data online in a secure and privacy-protecting manner [62–64]. Next to that, publication of curated, well-documented datasets is important as well. This is realised, among others, by offering research datasets that were used for publications in dedicated data-publishing repositories. Something that is compulsory in the genomic research area.

To this end, we have introduced the open-access CancerData biomedical data repository, which offers full datasets (e.g. in DICOM RT format) as well as supplemental data of accompanying publications. Furthermore, we introduced persistent identifiers using the Digital Object Identifier (DOI) coding system for the collected datasets. Referencing of individual DICOM objects using unique object identifiers (OID) is currently not supported, although a MAASTRO specific OID is assigned to every DICOM file, using a publicly registered IANA root id.

The allocation of DOI's allows researchers to easily cite the use of and allocate credit for the data, which stimulates scientific reuse and should result in a fairer assessment of academic performance [65,66]. A more fine-grained form of semantic-interoperable publishing of data is by means of *nano-publications* that use richly annotated RDF concepts and should lead to reduction of inconsistency, ambiguity and redundancy of data [67–70].

Changing clinical decisions

Rapid Learning is changing clinical decisions; by process and consequently by content, e.g. in the way they are made, as well as for the factual decisions themselves. The amount of information that is available nowadays and the way it is presented in the daily clinical workflow is getting beyond comprehension [21]. For instance, genomics or the imaging counterpart of Radiomics is generating huge amounts of raw data, which are impossible to interpret without tooling that converts this into clinical reasoning. Furthermore, it is known that it is impossible to keep up with recent literature and in the meantime the lifetime of systematic reviews is about 5.5 years, which likely means that derived clinical guidelines and protocols do not keep in pace with rapid technological changes [71,72].

The knowledge-based approach that was described here can support in solving this issue. Data-mining applications can process the large amounts of biomedical information and learn complication prediction models that are increasingly suited to offer personalised outcome predictions. This will subsequently stimulate shared decision making with well-informed patients [73].

One example hereof is the prototype of Treatment Choice (treatmentchoice.org), which is a user-friendly and interactive decision aid tool that helps lung cancer patients with the choice between chemotherapy and radiotherapy [74,75]. Not only is this a valuable service for patients, in a more sophisticated version this would be invaluable for physicians as well for the use in multi-disciplinary boards where there are similar choices to be made. Providing multimodality, multifactorial decision aids that quickly process lab reports, actual used medications and a full treatment history would ease physicians of a burden, tremendously. Unfortunately, fully clinically applicable decision support systems are not current practice yet. Large-scale programmes such as the Watson Health project will hopefully soon deliver.

The efficiency paradox

The field of knowledge engineering in oncology is relatively new and it has the ability to offer real solutions to facilitate model-based decision-making. This has the potential to improve the quality of treatments. However, the quest for effective, yet efficient therapy finds itself in a paradox where the optimal process is only found by an increasing effort to measure this process and its outcome.

Great care should be taken when introducing new medical informatics technology into a busy clinical environment [76]. The treatment process is under

constant pressure with the increasing need of electronic registration and external monitoring of process efficiency. For instance, the introduction of decision-support systems or information systems to stimulate incorporation of patient preferences into clinical practice needs proper facilitation for both patients as well as the medical staff.

With this rapidly increasing amount of information that needs to be generated as well as processed by professionals, a disturbing mismatch can occur between privacy management and efficient clinical use [77]. Health informatics vendors should better focus on user-centred design principles and human factors theory in their user interfaces (e.g. ISO 9241-210) and use well-structured data storage definitions for efficient reuse [78–80].

Conclusions

Access to ample quantities of data while having good quality is a huge challenge, especially in healthcare. Many centres lack resources or do not focus to archive data for reuse and research. However, it should be manifest that it is essential to collectively invest in secure but easy access to good quality medical data. It is the author's opinion that there is an obligation of the radiotherapy (research) community to share the information, which is mainly gathered under public-funded projects, for further investigation by colleagues and subsequent feedback to society through free open-access data repositories and open-source tools, such as CancerData or PredictCancer.

Clinical decision-making by *in silico* comparison of expected dose levels and subsequent toxicities, using multifactorial prediction models is introduced in the Netherlands for proton therapy. We added a third layer to differentiate between clinical and cost-effectiveness. It is expected that this approach will be considered for other rapidly advancing technical treatment options as well. It is a sensible way to choose the right treatment for the right patient, keeping a national or even global healthcare budget into account. Next, it is essential to include patient preferences for further improvement of the method.

The discipline of Clinical Data Science has many challenges to conquer but also a vast arsenal of solutions to offer for the cause of individualising healthcare. The Rapid Learning approach will provide more and faster feedback from healthcare quality improvement interventions. It is crucial for the validation of this approach to continuously capture the data that is used and generated while looping through a PDSA process as depicted in Figure 8.
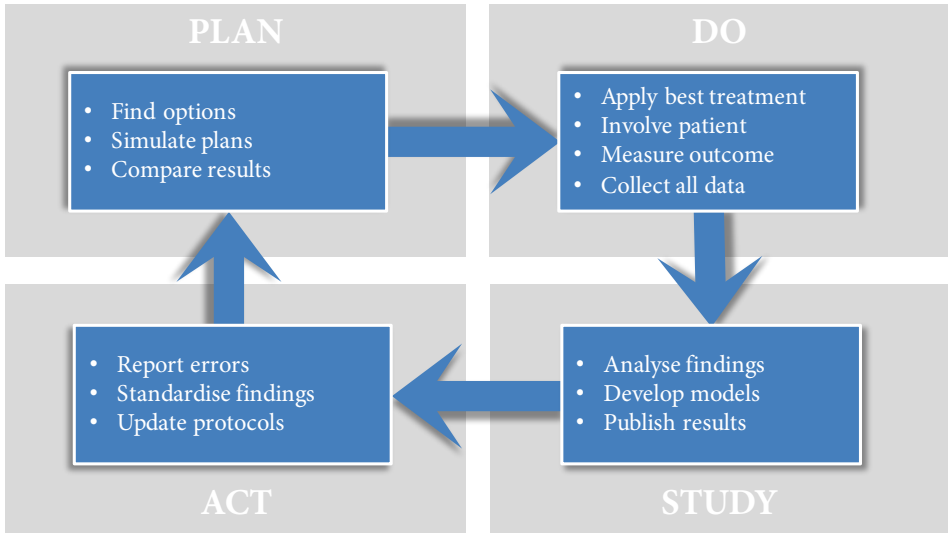
Figure 8 Rapid Learning PDSA cycle to continuously improve model-based radiotherapy

References

1. Lambin P, Petit SF, Aerts HJWL, van Elmpt WJC, Oberije CJG, Starmans MHW, *et al.* The ESTRO Breur Lecture 2009. From population to voxel-based radiotherapy: Exploiting intra-tumour and intra-organ heterogeneity for advanced treatment of non-small cell lung cancer. *Radiother Oncol* 2010;96(2):145–152. doi:10.1016/j.radonc.2010.07.001

2. Van Elmpt W, McDermott L, Nijsten S, Wendling M, Lambin P, Mijnheer B. A literature review of electronic portal imaging for radiotherapy dosimetry. *Radiother Oncol* 2008;88(3):289–309. doi:10.1016/j.radonc.2008.07.008

3. Nijsten SMJJG, Mijnheer BJ, Dekker ALAJ, Lambin P, Minken AWH. Routine individualised patient dosimetry using electronic portal imaging devices. *Radiother Oncol* 2007;83(1):65–75. doi:10.1016/j.radonc.2007.03.003

4. Van Elmpt W, Nijsten S, Mijnheer B, Dekker A, Lambin P. The next step in patient-specific QA: 3D dose verification of conformal and intensity-modulated RT based on EPID dosimetry and Monte Carlo dose calculations. *Radiother Oncol* 2008;86(1):86–92. doi:10.1016/j.radonc.2007.11.007

5. Van Elmpt WJC. 3D dose verification for advanced radiotherapy. Maastricht University, 2009

6. Nijsten SMJJG, Van Elmpt WJC, Mijnheer BJ, Minken AWH, Persoon LCGG, Lambin P, *et al.* Prediction of DVH parameter changes due to setup errors for breast cancer treatment based on 2D portal dosimetry. *Med Phys* 2009;36(1):83. doi:10.1118/1.3026660

7. Goldberg P, Conti RM. Problems With Public Reporting of Cancer Quality Outcomes Data. *J Oncol Pract* 2014;10(3):215–218. doi:10.1200/JOP.2014.001405

8. Bylicki O, Gan HK, Joly F, Maillet D, You B, Péron J. Poor patient-reported outcomes reporting according to CONSORT guidelines in randomized clinical trials evaluating systemic cancer therapy. *Ann Oncol Off J Eur Soc Med Oncol ESMO* 2015;26(1):231–237. doi:10.1093/annonc/mdu489

9. Faithfull S, Lemanska A, Chen T. Patient-reported Outcome Measures in Radiotherapy: Clinical Advances and Research Opportunities in Measurement for Survivorship. *Clin Oncol R Coll Radiol G B* 2015;27(11):679–685. doi:10.1016/j.clon.2015.07.007

10. Gilbert A, Ziegler L, Martland M, Davidson S, Efficace F, Sebag-Montefiore D, *et al.* Systematic Review of Radiation Therapy Toxicity Reporting in Randomized Controlled Trials of Rectal Cancer: A Comparison of Patient-Reported Outcomes and Clinician Toxicity Reporting. *Int J Radiat Oncol Biol Phys* 2015;92(3):555–567. doi:10.1016/j.ijrobp.2015.02.021

11. Campbell D. NHS failed to collect data on cancer treatment outcome. *The Guardian* 2015. www.theguardian.com/society/2015/sep/17/nhs-failed-collect-data-1bn-cancer-treatment-outcome-rejected-drugs-health [accessed November 8, 2015]

12. Morse A, NHS England, Great Britain, Department of Health, Great Britain, National Audit Office. *Investigation into the Cancer Drugs Fund: report.* 2015

13. Jenkins V. Discussing randomised clinical trials of cancer therapy: evaluation of a Cancer Research UK training programme. *BMJ* 2005;330(7488):400–0. doi:10.1136/bmj.38366.562685.8F

14. Grand MM, O'Brien PC. Obstacles to participation in randomised cancer clinical trials: a systematic review of the literature. *J Med Imaging Radiat Oncol* 2012;56(1):31–39. doi:10.1111/j.1754-9485.2011.02337.x

15. Murthy VH, Krumholz HM, Gross CP. Participation in cancer clinical trials: race-, sex-, and age-based disparities. *J Am Med Assoc* 2004;291(22):2720–2726. doi:10.1001/jama.291.22.2720

16. Movsas B, Moughan J, Owen J, Coia LR, Zelefsky MJ, Hanks G, *et al.* Who enrolls onto clinical oncology trials? A radiation Patterns Of Care Study analysis. *Int J Radiat Oncol Biol Phys* 2007;68(4):1145–1150. doi:10.1016/j.ijrobp.2007.01.051

17. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. *J R Soc Med* 2011;104(12):510–520. doi:10.1258/jrsm.2011.110180

18. Grutters JPC, Abrams KR, de Ruysscher D, Pijls-Johannesma M, Peters HJM, Beutner E, *et al.* When to wait for more evidence? Real options analysis in proton therapy. *The Oncologist* 2011;16(12):1752–1761. doi:10.1634/theoncologist.2011-0029

19. Sheehan M, Timlin C, Peach K, Binik A, Puthenparampil W, Lodge M, *et al.* Position statement on ethics, equipoise and research on charged particle radiation therapy. *J Med Ethics* 2013:medethics – 2012–101290. doi:10.1136/medethics-2012-101290

20. Etheredge LM. A Rapid-Learning Health System. *Health Aff (Millwood)* 2007;26(2):w107–w118. doi:10.1377/hlthaff.26.2.w107

21. Abernethy AP, Etheredge LM, Ganz PA, Wallace P, German RR, Neti C, *et al.* Rapid-Learning System for Cancer Care. *J Clin Oncol* 2010;28(27):4268–4274. doi:10.1200/JCO.2010.28.5478

22. Friedman CP, Wong AK, Blumenthal D. Achieving a Nationwide Learning Health System. *Sci Transl Med* 2010;2(57):57cm29–cm57cm29. doi:10.1126/scitranslmed.3001456

23. Moen RD, Norman CL. *Circling Back: clearing up myths about the Deming cycle and seeing how it keeps evolving.* American Society for Quality; 2010

24. Moen R. Foundation and History of the PDSA Cycle. www.deming.org/sites/default/files/pdf/2015/PDSA_History_Ron_Moen.pdf

25. Taylor MJ, McNicholas C, Nicolay C, Darzi A, Bell D, Reed JE. Systematic review of the application of the plan–do–study–act method to improve quality in healthcare. *BMJ Qual Saf* 2013:bmjqs – 2013–001862. doi:10.1136/bmjqs-2013-001862

26. Langley GJ, Moen R, Nolan K, Nolan T, Norman C, Provost L, editors. *The improvement guide: a practical approach to enhancing organizational performance.* 2nd ed. San Francisco: Jossey-Bass; 2009

27. Pareto V. *Manuale di economia politica.* Società Editrice Libraria; 1906

28. Hoffmann AL, Siem AYD, Hertog D den, Kaanders JHAM, Huizenga H. Derivative-free generation and interpolation of convex Pareto optimal IMRT plans. *Phys Med Biol* 2006;51(24):6349–6369. doi:10.1088/0031-9155/51/24/005

29. Thieke C, Küfer KH, Monz M, Scherrer A, Alonso F, Oelfke U, *et al.* A new concept for interactive radiotherapy planning with multicriteria optimization: first clinical evaluation. *Radiother Oncol* 2007;85(2):292–298. doi:10.1016/j.radonc.2007.06.020

30. Nelson EC, Eftimovska E, Lind C, Hager A, Wasson JH, Lindblad S. Patient reported outcome measures in practice. *The BMJ* 2015;350:g7818. doi:10.1136/bmj.g7818

31. European Commission. *Directive 2001/20/EC of the European parliament and of the Council of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use.* 2001

32. European Commission. *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.* 2001

33. Hoey R. The EU Clinical Trials Directive: 3 years on. *The Lancet* 2007;369(9575):1777–1778. doi:10.1016/S0140-6736(07)60797-1

34. Eurocan plus report: feasibility study for coordination of national cancer research activities. *Ecancermedicalscience* 2008;2. doi:10.3332/ecancer.2011.84

35. Sullivan R, Peppercorn J, Sikora K, Zalcberg J, Meropol NJ, Amir E, *et al.* Delivering affordable cancer care in high-income countries. *Lancet Oncol* 2011;12(10):933–980. doi:10.1016/S1470-2045(11)70141-3

36. Dekker A, Vinod S, Holloway L, Oberije C, George A, Goozee G, *et al.* Rapid learning in practice: a lung cancer survival decision support system in routine patient care data. *Radiother Oncol* 2014;113(1):47–53. doi:10.1016/j.radonc.2014.08.013

37. Damiani A, Gatta R, Dinapoli N, Meldolesi E, Jochems A, Deist T, *et al.* A New Model to Solve Privacy Issues in Multi-centric Clinical Studies: Problems Solved and New Challenges. *Int J Radiat Oncol* 2015;93(3, Supplement):E385–E386. doi:10.1016/j.ijrobp.2015.07.1530

38. National Platform of Proton therapy (LPPT). *Consensus document voor selectie van patiënten met een model-based indicatie voor protonen therapie (Unpublished internal document).* 2015

39. Lambin P, Zindler J, Vanneste B, van de Voorde L, Jacobs M, Eekers D, *et al.* Modern clinical research: How rapid learning health care and cohort multiple randomised clinical trials complement traditional

evidence based medicine. *Acta Oncol Stockh Swed* 2015;54(9):1289–1300. doi:10.3109/0284186X.2015.1062136

40. Langendijk JA, Lambin P, De Ruysscher D, Widder J, Bos M, Verheij M. Selection of patients for radiotherapy with protons aiming at reduction of side effects: the model-based approach. *Radiother Oncol* 2013;107(3):267–273. doi:10.1016/j.radonc.2013.05.007

41. Widder J, van der Schaaf A, Lambin P, Marijnen CAM, Pignol JP, Rasch CR, *et al.* The quest for evidence for proton therapy: the model-based approach and precision medicine. *Int J Radiat Oncol* 2015. doi:10.1016/j.ijrobp.2015.10.004

42. Council for Health and Society (RVS). *Sensible and sustainable care.* 2006

43. Vanneste BGL, Pijls-Johannesma M, Van De Voorde L, van Lin EN, van de Beek K, van Loon J, *et al.* Spacers in radiotherapy treatment of prostate cancer: is reduction of toxicity cost-effective? *Radiother Oncol* 2015;114(2):276–281. doi:10.1016/j.radonc.2015.01.005

44. Office of the National Coordinator for Health Information Technology (ONC). *Report on Health Information Blocking.* Washington: Office of the National Coordinator for Health Information Technology (ONC); 2015

45. Oberije C, Troost E, Eble M, Paul Bulens, Philippe Coucke, Katrien De Jaeger, *et al.* EuroCAT Umbrella Protocol for NSCLC. *CancerData* 2015. doi:10.17195/candat.2013.08.1

46. Langendijk JA, Lambin P. Standard Follow Up Program For Head And Neck Cancer Patients. *CancerData* 2012. doi:10.17195/candat.2012.01.1

47. Yue Cao. Advanced MR Imaging Protocol for Glioblastoma. *CancerData* 2016. doi:10.17195/candat.2016.03.1

48. Meldolesi E, van Soest J, Dinapoli N, Dekker A, Damiani A, Gambacorta MA, *et al.* An umbrella protocol for standardized data collection (SDC) in rectal cancer: a prospective uniform naming and procedure convention to support personalized medicine. *Radiother Oncol* 2014;112(1):59–62. doi:10.1016/j.radonc.2014.04.008

49. Santanam L, Hurkmans C, Mutic S, van Vliet-Vroegindeweij C, Brame S, Straube W, *et al.* Standardizing Naming Conventions in Radiation Oncology. *Int J Radiat Oncol* 2012;83(4):1344–1349. doi:10.1016/j.ijrobp.2011.09.054

50. Miller AA. A Rational Informatics-enabled approach to Standardised Nomenclature of Contours and Volumes in Radiation Oncology Planning. *J Radiat Oncol Inform* 2014;6(1):53–97

51. WHO. WHO | International Classification of Diseases (ICD). *WHO* 1994. www.who.int/classifications/icd/en [accessed November 8, 2015]

52. LOINC. Logical Observation Identifiers Names and Codes. loinc.org [accessed November 8, 2015]

53. NCI. NCI Thesaurus 1994. ncit.nci.nih.gov/ncitbrowser [accessed November 8, 2015]

54. IHTSDO. International Health Terminology Standards Development Organization. SNOMED Clinical Terms (SNOMED CT). www.ihtsdo.org/snomed-ct [accessed November 8, 2015]

55. EUROCRAN. Eurocan plus report: feasibility study for coordination of national cancer research activities. *Ecancermedicalscience* 2008;2:84. doi:10.3332/ecancer.2011.84

56. Tello A, Cruz AL, Saquicela V, Espinoza M, Vidal ME. RDF-ization of DICOM Medical Images towards Linked Health Data Cloud. In: Braidot A, Hadad A, editors. *VI Latin American Congress on Biomedical Engineering CLAIB 2014, Paraná, Argentina 29, 30 & 31 October 2014,* Springer International Publishing; 2015

57. Van Soest J, Lustberg T, Grittner D, Marshall MS, Persoon L, Nijsten B, *et al.* Towards a semantic PACS: Using Semantic Web technology to represent imaging data. *Stud Health Technol Inform* 2014:166–170. doi:10.3233/978-1-61499-432-9-166

58. Miller A. The future of health care could be elementary with Watson. *CMAJ Can Med Assoc J* 2013;185(9):E367–E368. doi:10.1503/cmaj.109-4442

59. Doyle-Lindrud S. Watson will see you now: a supercomputer to help clinicians make informed treatment decisions. *Clin J Oncol Nurs* 2015;19(1):31–32. doi:10.1188/15.CJON.31-32

60. Ratner M. IBM's Watson Group signs up genomics partners. *Nat Biotechnol* 2015;33(1):10–11. doi:10.1038/nbt0115-10

61. Oncologists Partner with Watson on Genomics. *Cancer Discov* 2015;5(8):788–788. doi:10.1158/2159-8290.CD-NB2015-090

62. Garde S, Knaup P, Hovenga E, Heard S. Towards semantic interoperability for electronic health records. *Methods Inf Med* 2007;46(3):332–343. doi:10.1160/ME5001

63. Klein A, Prokosch HU, Muller M, Ganslandt T. Experiences with an interoperable data acquisition platform for multi-centric research networks based on HL7 CDA. *Methods Inf Med* 2007;46(5):580–585. doi:10.1160/ME9060

64. Ibrahim A, Bucur A, Perez-Rey D, Alonso E, de Hoog M, Dekker A, *et al.* Case Study for Integration of an Oncology Clinical Site in a Semantic Interoperability Solution based on HL7 v3 and SNOMED-CT: Data Transformation Needs. *AMIA Jt Summits Transl Sci Proc AMIA Summit Transl Sci* 2015;2015:71

65. Credit where credit is due. *Nature* 2009;462(7275):825–825. doi:10.1038/462825a

66. Lane J. Let's make science metrics more scientific. *Nature* 2010;464(7288):488–489. doi:10.1038/464488a

67. Mons B, Velterop J. Nano-publication in the e-science era, *Workshop on Semantic Web Applications in Scientific Discourse (SWASD 2009)*. 2009

68. Groth P, Gibson A, Velterop J. The anatomy of a nanopublication. *Inf Serv Use* 2010(1-2):51–56. doi:10.3233/ISU-2010-0613

69. Patrinos GP, Cooper DN, van Mulligen E, Gkantouna V, Tzimas G, Tatum Z, *et al.* Microattribution and nanopublication as means to incentivize the placement of human genome variation data into the public domain. *Hum Mutat* 2012;33(11):1503–1512. doi:10.1002/humu.22144

70. Christine C, Oliver K, Pascale G, Lydie L, Barend M, Amos B. Converting neXtProt into Linked Data and nanopublications. *Semantic Web* 2015(2):147–153. doi:10.3233/SW-140149

71. Glasziou PP, Sawicki PT, Prasad K, Montori VM. Not a Medical Course, but a Life Course: *Acad Med* 2011;86(11):e4. doi:10.1097/ACM.0b013e3182320ec9

72. The half-life of facts. *The Economist* 2012

73. Oberije C, Nalbantov G, Dekker A, Boersma L, Borger J, Reymen B, *et al.* A prospective study comparing the predictions of doctors versus models for treatment outcome of lung cancer patients: a step toward individualized care and shared decision making. *Radiother Oncol* 2014;112(1):37–43. doi:10.1016/j.radonc.2014.04.012

74. Stacey D, Légaré F, Col NF, Bennett CL, Barry MJ, Eden KB, *et al.* Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev* 2014;1:CD001431. doi:10.1002/14651858.CD001431.pub4

75. Berlanga AJ, Roelofs E, Troost E, Braeken A, Nalbantov G, Lambin P. EP-1378: Considering patient preferences to choose the best individual treatment, the example of a treatment choice tool. *Radiother Oncol* 2014;111:S114. doi:10.1016/S0167-8140(15)31496-1

76. Garcia R, Nyström H, Fiorino C, Thwaites D. Does a too risk-averse approach to the implementation of new radiotherapy technologies delay their clinical use? *Br J Radiol* 2015;88(1051):20150124. doi:10.1259/bjr.20150124

77. Eikey EV, Murphy AR, Reddy MC, Xu H. Designing for privacy management in hospitals: Understanding the gap between user activities and IT staff's understandings. *Int J Med Inf* 2015;84(12):1065–1075. doi:10.1016/j.ijmedinf.2015.09.006

78. Carayon P, Wetterneck TB, Rivera-Rodriguez AJ, Hundt AS, Hoonakker P, Holden R, *et al.* Human factors systems approach to healthcare quality and patient safety. *Appl Ergon* 2014;45(1):14–25. doi:10.1016/j.apergo.2013.04.023

79. Ratwani RM, Fairbanks RJ, Hettinger AZ, Benda NC. Electronic health record usability: analysis of the user-centered design processes of eleven electronic health record vendors. *J Am Med Inform Assoc JA-MIA* 2015. doi:10.1093/jamia/ocv050

80. Stanziola E, Uznayo MQ, Ortiz JM, Simón M, Otero C, Campos F, *et al.* User-Centered Design of Health Care Software Development: Towards a Cultural Change. *Stud Health Technol Inform* 2015;216:368–371. doi:10.3233/978-1-61499-564-7-368

Valorisation Addendum

Relevance of this work

The research activities that are presented in this dissertation primarily concentrate on two intertwined topics: 1) to investigate the need and possible solutions for international research data infrastructures to collaboratively capture, process and share patient information ranging from every day clinical data records to highly curated and enriched clinical trial databases. And 2) to investigate methods of application of data-driven prediction models to improve the individualised prescription of cancer therapy options, in particular when considering costly treatment modalities such as charged particle therapy.

Both areas of research are cornerstones of the concept of Rapid-Learning Healthcare (RLHC), as introduced in Chapter 2 and contribute to much needed solutions to bring research findings into clinical practice more rapidly. Furthermore, with the current economic healthcare climate we are in and the aging population, we face challenging times where the request for clinical effectiveness increasingly becomes a quest for cost-effectiveness. The application of the RLHC framework helps to address this and contributes to a better healthcare system in three ways:

- To clinicians new tooling is provided to take decisions based on larger amounts of patient-specific clinical, biological and social information. This enables them to choose for more personalised treatment options, while keeping the scarce resources and limited healthcare budgets into account.

- For patients there is the benefit of large quantities of historical datasets that are connected into a globe-spanning learning system, which facilitates the matching of their individual disease profiles to promising new treatments or clinical trials.

- Furthermore, due to consistent documentation of new complication prediction models, treatment protocols and patient outcomes, healthcare managers and policymakers can constantly be informed with up-to-date, high-quality data of the treatments options and their results. This helps to improve the targeting of healthcare budgets and investments to enable the right treatments for the right patients.

Although, RLHC was introduced around five years ago, we still experience little of it in daily clinical practice. Unfortunately, it is not unusual in healthcare to take more than ten years for innovative ideas to translate into clinical practice. The research as described in this dissertation, however, assists in realising

RLHC by offering several pioneering concepts and components [1] that can be put it into use through individual implementation or via commercialisation.

Dissemination and utilisation

Research data infrastructures

As described in Chapter 4, we created a unique data exchange platform to share information from electronic patient records and full image datasets in a secure de-identified manner between international hospitals. The system was built with open-source and freely available software and explained in detail to allow for similar initiatives to be undertaken. Such a one-to-one solution is very valuable for initial, small-scale investigations, however a wider scope is warranted. The large amount and diversity of clinical data that is needed to keep answering increasingly complex questions is demanding the combination of many more institutions to share their information from locked-away data silos.

One such a multicentric research data infrastructure is the BioMedical Imaging Archive (BMIA, bmia.nl), which we set up when leading the medical imaging work package in the Translational Research IT (TraIT) project of the Center for Translational Molecular Medicine (CTMM, now part of the Lygature foundation). BMIA is a centralised national archive based on the open-source National Biomedical Imaging Archive (NBIA) software offered by the NCI's National Cancer Informatics Program. The platform offers the sustained storage of previously collected research data or from new multicentric clinical trials, for instance. CTMM is offering the BMIA service as part of TraIT's sustainable infrastructure to the Dutch medical community in a 'freemium' manner allowing small projects to use the services for free.

The MISTIR (mistir.info) research platform for multicentric *in silico* trials (Chapter 5), which was created with open-source software as well, was specifically set up to run studies that compare current state-of-the-art photon therapy to promising charged particle therapies. The Radiation Oncology Collaborative Comparison (ROCOCO) consortium is running multiple simulated clinical trials that require the exchange of de-identified clinical cases in a safe and efficient manner. Although the system and trial data are not publicly available via MISTIR, we are preparing the datasets to be released via the CancerData portal.

The CancerData portal (cancerdata.org) is an open, online platform that offers permanent public access to supplementary datasets of publications in the

field of radiation oncology [2]. Various data formats are offered, ranging from full DICOM (RT) datasets via a locally installed NBIA repository, to documents, spreadsheets and graphs. CancerData also offers several standardised research protocols, including material transfer agreements, scoring of side effects, quality of life questionnaires and patient information sheets in multiple languages, among others. By collectively using these standardised data collection protocols, the data quality is significantly improved and it is much easier to gain new insights and learn prediction models from pooled multiple databases. Sharing that information with other institutions enables them to continue with subsequent research and to use that information to verify or improve previous findings or use the data to verify their own prediction models, for instance.

Unlocking this potential knowledge by reusing the medical records stored within hospital walls via public repositories is a huge benefit to the oncology community. However, sharing sensitive patient data is a big challenge as personal information needs to be removed from the data while maintaining the value within the records to still be useful. While this can be established with proven open-source tooling, it is undoable in the long run to replicate all relevant clinical data into research archives (may they be public or not). Not only because the data generated before and during a treatment is rapidly expanding, but also because we are able to find important information hidden within currently –mistakenly interpreted– irrelevant day-to-day clinical data.

That is where the work as described in Chapters 3, 4 and §8.3 comes in. This research, which was performed in the euregional EuroCAT and many subsequent international projects, has led to a worldwide medical infrastructure with over 20 nodes, which enables a federated learning network in which medical data can be processed in a privacy-preserving way to study disease and patient characteristics in relation to treatment options and patient outcomes. Each institution in the network becomes a node that has dedicated software installed, which translates the clinical patient data into anonymous codes and statistical information. This unique, powerful data liberation technique enables the pooling of many global datasets to increase prediction power and allow independent validation of discovered relations. The federated learning network methodology and various software components that have been developed during the research are co-developed with Siemens, Varian Medical Systems, Philips and multiple SME showcasing the intense interest also from commercial entities in this approach.

Decision support

As described in Chapter 7, we developed a prototype for a proton decision support system (PRODECIS, prodecis.nl), based on the data and experience from the ROCOCO trials. The model-based approach that has been introduced in the Netherlands for proton therapy of lung, head-and-neck, prostate and breast cancer requires Dutch proton therapy centres to show an estimated benefit of proton therapy before treatments are being reimbursed. This estimation is given by comparing the prediction of complications that can occur after applying photon- or proton-based radiotherapy. Only when the expected benefit is exceeding a nationally defined threshold, the centre will be reimbursed. The system has been licensed to ptTheragnostic, which is a spin-out of MAASTRO CLINIC. Initially, it is being developed for the European market but is expected to be available worldwide in the coming year.

The model-based approach for proton therapy is a unique method that was developed in the last couple of years. It is based on clinical prediction models, however, which are not new at all, although clinical application is still limited. To stimulate the research, development and actual use of decision support systems, we publish prediction models online on PredictCancer (predictcancer.org). For the PRODECIS system, we used one such a published head-and-neck complication prediction model that includes a cost-effectiveness model comparing photon and proton therapy.

Open Data, Open Knowledge

As indicated previously, we actively use and support the open-source or open-access philosophy. In general, we adhere as much as possible to the "open knowledge" principles given by the Open Knowledge Foundation (OKFN, okfn.org):

> *"'Open knowledge' is any content, information or data that people are free to use, re-use and redistribute — without any legal, technological or social restriction."*

Therefore, the data that are the basis of derived knowledge should be made available for reuse whenever possible. Or as given by the OKFN:

> *"Everyone should have access to publicly-funded research in all areas, including the latest science, medicine, and culture — both scholarly papers*

and research data. Essential information such as the results of medical trials, successful or otherwise, should be openly available to all."

Doing so will 1) enable transparency such that any individual is granted insight in the information and is given a means to check previous findings, for instance, 2) release social and possibly commercial value, which is exactly the intension of 'valorisation' and 3) stimulate participation and engagement by other researchers, patients and policy makers, for instance.

Another data-sharing stimulus focussed on internationally supported data stewardship, has been proposed by the Dutch Techcentre for Life Sciences (DTL) via the "FAIR Data" principles [3], in which data are:

- "Findable: easy to find by both humans and computer systems and based on mandatory description of the metadata that allow the discovery of interesting datasets,

- Accessible: stored for long term such that they can be easily accessed and/or downloaded with well-defined license and access conditions (Open Access when possible), whether at the level of metadata, or at the level of the actual data content,

- Interoperable: ready to be combined with other datasets by humans as well as computer systems and

- Reusable: ready to be used for future research and to be processed further using computational methods."

To this end, the data should be open both legally, that is under a license that permits free access, reuse, modification and redistribution of the data, and technically, being easy to access the complete datasets in 'bulk' and machine-readable, e.g. as 'linked data'. Adhering to these principles, we offer the three basic components of our research activities to the public:

- Publications are published open access whenever possible. When the publication is copyrighted by a journal, it is often possible to offer a Final Author Version (the authors' accepted version of their manuscript), referring to the published version.

- Prediction models are freely accessible on PredictCancer (predictcancer.org). It is also possible for external parties to submit their models to the site for publication.

- Data that were derived from our research are published on CancerData (cancerdata.org). To offer the material in a machine-readable manner and

in the meantime boost the online presence of the publications and their datasets, we generate persistent links to the datasets via the Document Object Identifier (DOI) mechanism. To further improve the level of visibility of the research and its data, we have CancerData being harvested for content by OpenDOAR, BioSharing and Re3Data indexing sites, among others.

Such open repositories provide a means to directly show appreciation for other peoples' work (e.g. by *liking* or *ranking*), which is a popular, contemporary social-network habit and a direct incentive instead of the slow-paced impact-factor-driven academic reward system.

We will continue working on both the future-proof, global federated learning network that will handle *big data* research across large disparate medical databases as well as centralised *linked data* repositories with open access to the outcome from that research.

References

1. Skripcak T, Belka C, Bosch W, Brink C, Brunner T, Budach V, *et al.* Creating a data exchange strategy for radiotherapy research: towards federated databases and anonymised public datasets. *Radiother Oncol* 2014;113(3):303–309. doi:10.1016/j.radonc.2014.10.001

2. re3data.org. CancerData.org 2014. doi:10.17616/R38W4R

3. Wilkinson MD, Dumontier M, Aalbersberg IjJ, Appleton G, Axton M, Baak A, *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 2016;3:160018. doi:10.1038/sdata.2016.18

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 beslishulpsystemen 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 protonentherapie.

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 protonentherapie de dosis op het normale weefsel kan verminderen bij gelijke of hogere dosis op de tumor. De deelnemende centra downloadden 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 protonentherapie 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 protonentherapie 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

LIST OF
PUBLICATIONS

- Walsh S, **Roelofs E**, Kuess P, Lambin P, Jones B, Georg D, Verhaegen F. A validated tumor control probability model based on a meta-analysis of low, intermediate, and high-risk prostate cancer patients treated by photon, proton, or carbon-ion radiotherapy. Med Phys. 2016 Feb;43[2]:734. doi:10.1118/1.4939260

- Cheng Q, **Roelofs E**, Ramaekers BLT, Eekers D, van Soest J, Lustberg T, Hendriks T, Hoebers F, van der Laan HP, Korevaar EW, Dekker A, Langendijk JA, Lambin P. Development and evaluation of an online three-level proton vs photon decision support prototype for head and neck cancer – Comparison of dose, toxicity and cost-effectiveness. Radioth Oncol. 2016 Feb;118[2]:281–5. doi:10.1016/j.radonc.2015.12.029

- van Soest JPA, Dekker ALAJ, **Roelofs E**, Nalbantov G. Application of Machine Learning for Multicenter Learning [Internet]. In: El Naqa I, Li R, Murphy JM, editors. Machine Learning in Radiation Oncology: Theory and Applications. Cham: Springer International Publishing; 2015. p. 71–97. doi:10.1007/978-3-319-18305-3_6

- Lambin P, Zindler J, Vanneste B, van de Voorde L, Jacobs M, Eekers D, Peerlings J, Reymen B, Larue RTHM, Deist TM, de Jong EEC, Even AJG, Berlanga AJ, **Roelofs E**, Cheng Q, Carvalho S, Leijenaar RTH, Zegers CML, van Limbergen E, Berbee M, van Elmpt W, Oberije C, Houben R, Dekker A, Boersma L, Verhaegen F, Bosmans G, Hoebers F, Smits K, Walsh S. Modern clinical research: How rapid learning health care and cohort multiple randomised clinical trials complement traditional evidence based medicine. Acta Oncol. 2015 Oct;54[9]:1289–300. doi:10.3109/0284186X.2015.1062136

- Wink KCJ, **Roelofs E**, Solberg T, Lin L, Simone CB, Jakobi A, Richter C, Lambin P, Troost EGC. Particle therapy for non-small cell lung tumors: where do we stand? A systematic review of the literature. Front. Oncol. 2014;4:292. doi:10.3389/fonc.2014.00292

- **Roelofs E**, Dekker A, Meldolesi E, van Stiphout RGPM, Valentini V, Lambin P. International data-sharing for radiotherapy research: An open-source based infrastructure for multicentric clinical data mining. Radioth Oncol. 2014 Feb;110[2]:370–4. doi:10.1016/j.radonc.2013.11.001

- Walsh S, **Roelofs E**, Verhaegen F. Comparing geometrical plan robustness and volatility of TCP for the ROCOCO photon prostate dataset. Physica Medica. 2013 Sep;29[5]:571. doi:10.1016/j.ejmp.2013.05.021

- **Roelofs E**, Persoon L, Nijsten S, Wiessler W, Dekker A, Lambin P. Benefits of a clinical data warehouse with data mining tools to collect data for a radiotherapy trial. Radiother Oncol. 2013 Jul;108[1]:174–9. doi:10.1016/j.radonc.2012.09.019

- Lambin P, van Stiphout RGPM, Starmans MHW, Rios-Velazquez E, Nalbantov G, Aerts HJWL, **Roelofs E**, van Elmpt W, Boutros PC, Granone P, Valentini V, Begg AC, De Ruysscher D, Dekker A. Predicting outcomes in radiation oncology--multifactorial decision support systems. Nat Rev Clin Oncol. 2013 Jan;10[1]:27–40. doi:10.1038/nrclinonc.2012.196

- Lambin P, **Roelofs E**, Reymen B, Velazquez ER, Buijsen J, Zegers CML, Carvalho S, Leijenaar RTH, Nalbantov G, Oberije C, Scott Marshall M, Hoebers F, Troost EGC, van Stiphout RGPM, van Elmpt W, van der Weijden T, Boersma L, Valentini V, Dekker A. 'Rapid Learning health care in oncology' - an approach towards decision support systems enabling customised radiotherapy'. Radiother Oncol. 2013 Oct;109[1]:159–64. doi:10.1016/j.radonc.2013.07.007

- Fontanarosa D, van der Laan HP, Witte M, Shakirin G, **Roelofs E**, Langendijk JA, Lambin P, van Herk M. An in silico comparison between margin-based and probabilistic target-planning approaches in head and neck cancer patients. Radiother Oncol. 2013 Dec;109[3]:430–6. doi:10.1016/j.radonc.2013.07.012

- **Roelofs E**, Engelsman M, Rasch C, Persoon L, Qamhiyeh S, de Ruysscher D, Verhaegen F, Pijls-Johannesma M, Lambin P, ROCOCO Consortium. Results of a multicentric in silico clinical trial (ROCOCO): comparing radiotherapy with photons and protons for non-small cell lung cancer. J Thorac Oncol. 2012 Jan;7[1]:165–76. doi:10.1097/JTO.0b013e31823529fc

- Levy MA, Freymann JB, Kirby JS, Fedorov A, Fennessy FM, Eschrich SA, Berglund AE, Fenstermacher DA, Tan Y, Guo X, Casavant TL, Brown BJ, Braun TA, Dekker A, **Roelofs E**, Mountz JM, Boada F, Laymon C, Oborski M, Rubin DL. Informatics methods to enable sharing of quantitative imaging research data. Magn Reson Imaging. 2012 Nov;30[9]:1249–56. doi:10.1016/j.mri.2012.04.007

- Lambin P, van Stiphout RGPM, Starmans MHW, Rios-Velazquez E, Nalbantov G, Aerts HJWL, **Roelofs E**, van Elmpt W, Boutros PC, Granone P, Valentini V, Begg AC, De Ruysscher D, Dekker A. Predicting outcomes in

radiation oncology—multifactorial decision support systems. Nat Rev Clin Oncol. 2012 Nov 20;10[1]:27–40. doi:10.1038/nrclinonc.2012.196

- **Roelofs E**, Persoon L, Qamhiyeh S, Verhaegen F, De Ruysscher D, Scholz M, Iancu G, Engelsman M, Rasch C, Zijp L, Meerleer GD, Coghe M, Langendijk J, Schilstra C, Pijls-Johannesma M, Lambin P. Design of and technical challenges involved in a framework for multicentric radiotherapy treatment planning studies. Radioth Oncol. 2010 Dec;97[3]:567–71. doi:10.1016/j.radonc.2010.08.009

- Musat E, **Roelofs E**, Bar-Deroma R, Fenton P, Gulyban A, Collette L, Stupp R, Weber DC, Bernard Davis J, Aird E, Baumert BG. Dummy run and conformity indices in the ongoing EORTC low-grade glioma trial 22033-26033: First evaluation of quality of radiotherapy planning. Radioth Oncol. 2010 May 1;95[2]:218–24. doi:10.1016/j.radonc.2010.03.005

- de Xivry JO, **Roelofs E**, Piccoli T, Lambin P, Goebel R, Di Salle F, Veninga T, Macq B, Baumert B. Evaluating the impact of radiotherapy on brain tumour patients using fMRI and DTI: a pilot study. Radioth Oncol. 2009 Aug 1;92:S160. doi:10.1016/S0167-8140(12)73006-2

- Eekers DBP, Laschet A, de Groot M, **Roelofs E**, Kester A, Delaere K, Lambin P, van Gils F, Nap M, ten Kate J. Why determine only the total prostate-specific antigen, if the free-to-total ratio contains the information? Ann. Clin. Biochem. 2008 May;45[Pt 3]:270–4. doi:10.1258/acb.2007.007014

DANKWOORD

Allereerst hartelijk dank aan de leescommissie en coronaleden voor jullie bereidheid tijd te steken in het beoordelen van mijn dissertatie en de verdediging daarvan.

Beste Philippe, dank dat je me de mogelijkheid gaf om toe te treden tot de wondere wereld van het translationele onderzoek. Met je visie en tomeloze energie is het soms moeilijk je bij te houden, maar het doet ons uiteindelijk altijd tot grote hoogte stijgen. We hebben ons met het ROCOCO-team voortdurend hard gemaakt voor protonentherapie, van niemandsland (Trilandis) tot in huis (Maastro). Het uur van de waarheid zal binnenkort slaan. Op naar de volgende uitdaging!

Beste André, dank je voor je steun. Je was als een mentor voor me en het doet me erg veel deugd om je nu als promotor toe te kunnen voegen. Met het zogenaamde *MIKE-team* hebben we menige digitale speeltuin bezocht en steeds een stukje verder kunnen helpen opbouwen. Dat we heel trots mogen zijn op de resultaten op het gebied van *Clinical Data Science* blijkt bij menig congres of uit de wereldwijde samenwerkingen. Het was een eer om je soms te mogen vervangen en ik wens je alle succes om de wereld te veroveren met het vernieuwende klinische semantische web. Ik heb er alle vertrouwen in!

Beste André Minken, een speciaal woord van dank voor jou. Ik ben nog steeds heel blij dat je me welkom hebt geheten binnen de radiotherapie en mijn potentie als klinisch fysicus hebt gezien. Ook Iaïn Bruinvis en Ben Mijnheer ben ik veel dank verschuldigd voor jullie bijdrage daarin.

Heel veel dank aan het MAASTRO ROCOCO-team; zonder jullie was de protonencomponent in deze promotie niet mogelijk geweest. In alfabetische volgorde: Ben, Daniëlle, Dirk, Esther B., Esther T., Frank H., Frank V., Jørgen, Judith, Krista, Lucas, Madelon, Marlies, Philippe, Seán, Sima, Yvonne.

Thanks to all other (international) ROCOCO-members. It is undoable to name all of you personally here. Thanks to your efforts, we were able to show that proton therapy is a powerful and effective addition to the radiotherapy arsenal. You helped making it real for us. Let's keep up the good work!

Also a big shout-out (I hope you hear it all the way in China) to Qing Cheng. It was a pleasure to supervise and work with you in the PRODECIS project. Thanks to your contributions, we have a working prototype of the proton decision support system.

Hartelijk dank aan alle collega's bij MAASTRO waar ik in meer of mindere mate mee heb samengewerkt, zoals I&S (Natascha, dank je voor je support om

dit werk allemaal te kunnen doen), FI, TiQC, SDT, KE (Johan in het bijzonder voor alle TraIT en CancerData hulp), KIT, DCM (met name Ruud, voor je statistisch significante bijdrage) en andere mogelijke afkortingen die ik nu aan jullie over laat om te bedenken, (onderzoeks)laboranten, secretariaten en iedereen die ik vergeten ben. Kortom: "ALL Medewerkers MAASTRO clinic" bedankt!

Beste ProTones, m'n *happy* muziekvrienden, een groot applaus voor jullie! Dank dat ik altijd bij jullie terecht kan. Soms om een gevoelige snaar te raken, soms gewoon om alles er even uit te gooien, of het nu in de kelder van MAASTRO of in Barcelona is… dank dat ik het allemaal mee mag maken. Wie weet waar we volgend jaar staan? Ik doe graag weer mee!

Lieve Mariet en Albert, dank dat jullie me hebben verwelkomd in jullie gezin en mij jullie dochter hebben toevertrouwd. Jullie hebben me als didacten steeds gestimuleerd om blijvend te investeren. Het klopt, je bent nooit te oud om te leren, en blijkbaar zelfs, te promoveren. En zoals oma Delden altijd zei: "*'t leert en leert en wordt niet meu, wie kriegt nog 'n gebrek an domme leu!*".

Lieve Corry, regelmatig tufte je in je Panda van Amsterdam naar het diepe zuiden om op de kinderen te passen, de halve klerenkast te vernaaien of te haasten naar het postkantoor als ik zonder paspoort in het buitenland zat en weer naar huis moest zien te komen. Dank je voor alle hulp! En Thom, je zou ervan genoten hebben vandaag. Je bent er in gedachten bij!

Een groot woord van dank aan Jacco en Gerben, m'n dierbare paranimfen, die me geholpen hebben met de voorbereidingen en het volgehouden hebben mij bij te staan vandaag. Beste Gerben, we hebben samen al heel veel avonturen meegemaakt. Of het nu crossen in de kar is, met een vers rijbewijs en stevige blues-rock een eindje 'rijden' door de Holterberg, koprollen met de auto in Frankrijk of met de voorruit op schoot in de camper. Been there, done that and we'll just keep on rolling! Dank je dat je er al die jaren nog steeds voor me bent en speciaal vandaag!

Lieve Jacco, grote broer, je organisatietalent kwam goed van pas in de aanloop naar vandaag; dank je wel!. Met een paar jaar extra, heb je al vele wegen vóór mij gelopen, zodat ik er vaak mooi achteraan kon. Ik weet niet of je het er destijds om deed met onze 'robbertjes' op de mat (en maar niet afkloppen…), maar het heeft me zeker geleerd om niet op te geven!

Farida, ma chère sœur, jouw vrolijke lach is een aanwinst voor de familie! Je smakelijke verhalen en dito kippenpootjes op de verjaardagen zijn altijd weer om van te genieten!

Lieve mama en papa, jullie hebben me altijd gesteund, ook al koos ik soms paden die niet altijd voor de hand lagen, zoals bijvoorbeeld het verruilen van pingelen op de Spaanse gitaar voor jengelen op een elektrische. Ik heb hele fijne herinneringen aan een gelukkige jeugd: van dappere, maar jammerlijk mislukte pogingen om me minder spekjes te laten snoepen uit het pakhuis (Nijverdalse Banket Centrale), tot genieten van het buitenleven op het erf in Schuilenburg (behalve dan dat grasmaaien!). Jullie hebben ons keer op keer geweldig geholpen; van sloop- en schilderwerk tot oppassen op de kinderen en nu onze lieve pup, Rocky (waar ik heel erg blij mee ben!), tuineren en ook weer mooi *ankeer'n* na de klus. Dank voor alles.

Lieve Karsten en Imke, m'n grote schatten. Jullie zijn zulke lieve, vrolijke kinderen en al zo zelfstandig; ik ben ontzettend trots op jullie! De manier waarop jullie thuis meehelpen, bijvoorbeeld als ik weer eens op pad moet, is super. Eindelijk heb ik weer meer tijd om met jullie op de trampoline te springen. Eens kijken wie is er het eerst op is!

Lest best: lieve Jelle, *heel mijn hart*. Dank je voor alles. Je bent m'n steun en toeverlaat en verdient veel meer lof en waardering dan ik in deze paar regels kan opschrijven. Zonder jou zorg en support was het me nooit gelukt dit tot een goed einde te brengen. Je hebt zoveel voor me gedaan; nu ben jij aan de beurt!

Copyrights

H.A. (Erik) Roelofs, commonly known as ER(ik), was born on April 23rd, 1971, in Almelo. He is a medical physicist and scientist with a history in system engineering, involving project management and product design in a holistic, *mechatronic* approach. This overarching profile is reflected in his work combining the domains of medical physics and informatics with the development of medical data exchange infrastructures and knowledge-driven decision support systems, among others.

He pioneered in the set-up of a Bio-Medical Imaging Archive of the CTMM-TraIT project (*bmia.nl*), which enables a national, centralised research infrastructure to collect and share biomedical images and data. Furthermore, he developed the MISTIR collaboration platform for secure data exchange (*mistir.info*), including quality assurance measures and analyses procedures, to enable international, multicentric *in silico* ROCOCO trials investigating the possible benefit of particle therapy.

He is involved in technology transfer activities via the ptTheragnostic spin-off company (*pttheragnostic.com*) of MAASTRO clinic that sets out to *enable the right treatment for the right patient*. He leads the development of PRODECIS, which is an online, multi-level decision support system that facilitates a rapid, model-based stratification, which is needed in the Netherlands for the assessment of proton therapy for breast, head and neck, lung and prostate cancer.

He continues his work at MAASTRO clinic, connecting medical physics and information management. His current efforts include the proper reuse of medical information in international settings, such as enabling data micro-publishing via online research information systems (*cancerdata.org*). Furthermore, he is leading the Medical Informatics team, which focusses on the safe introduction and sustained use of innovative medical software in a busy, clinical environment. He is happily married to Daniëlle Eekers and a proud father of Karsten and Imke.

ORCiD