Learning methods in radiation oncology

‘Rapid Learning health care in oncology’ – An approach towards decision support systems enabling customised radiotherapy

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

a Department of Radiation Oncology (MAASTRO), GROW – School for Oncology and Developmental Biology; b Department of General Practice, CAPHRI – School for Public Health and Primary Care, Maastricht University, The Netherlands; c Department of Radiotherapy, Università Cattolica S. Cuore, Rome, Italy

A R T I C L E   I N F O

Article history:
Received 27 May 2013
Received in revised form 30 June 2013
Accepted 16 July 2013
Available online 28 August 2013

Keywords:
Radiotherapy
Decision support system (DSS)
Rapid Learning
Cancer
Tailored radiation treatment

A B S T R A C T

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.

© 2013 The Authors. Published by Elsevier Ireland Ltd. Open access under CC BY-NC-ND license. Radiotherapy and Oncology 109 (2013) 159–164

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.

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 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 http://www.predictcancer.org, breast cancer on http://research.nki.nl/ibr/ and glioblastoma on http://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.

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 multifactorial 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].

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 approaches; (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.optiongrid.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

We would like to acknowledge the financial support of the EU IM programme (QuIC-ConCePT), the CTMM framework (AIRFORCE project), the EU 6th and 7th framework programme (Metoxia, Artforce, Eureca), Interreg (www.eurocat.info) and the Dutch Cancer Society (KWF UIM 2011-5202, KWF UOM 2009-4454).

References

