Organism-level models: When mechanisms and statistics fail us

M H Phillips¹, J Meyer¹, W P Smith² and J K Rockhill¹

¹ Department of Radiation Oncology, University of Washington, Seattle, WA, USA
² Stratton VA Medical Center, Albany, NY, USA

E-mail: markp@u.washington.edu

Abstract. Purpose: To describe the unique characteristics of models that represent the entire course of radiation therapy at the organism level and to highlight the uses to which such models can be put. Methods: At the level of an organism, traditional model-building runs into severe difficulties. We do not have sufficient knowledge to devise a complete biochemistry-based model. Statistical model-building fails due to the vast number of variables and the inability to control many of them in any meaningful way. Finally, building surrogate models, such as animal-based models, can result in excluding some of the most critical variables. Bayesian probabilistic models (Bayesian networks) provide a useful alternative that have the advantages of being mathematically rigorous, incorporating the knowledge that we do have, and being practical. Results: Bayesian networks representing radiation therapy pathways for prostate cancer and head & neck cancer were used to highlight the important aspects of such models and some techniques of model-building. A more specific model representing the treatment of occult lymph nodes in head & neck cancer were provided as an example of how such a model can inform clinical decisions. A model of the possible role of PET imaging in brain cancer was used to illustrate the means by which clinical trials can be modelled in order to come up with a trial design that will have meaningful outcomes. Conclusions: Probabilistic models are currently the most useful approach to representing the entire therapy outcome process.

1. Introduction

Models are of interest in medical science because they promote our understanding of fundamental processes. Just as importantly, however, they allow us to predict whether it be the type of disease, its future progression or its response to a specific intervention. A more complete model treats more than just the disease or intervention, but also its interaction with all aspects of a person’s health. In radiation therapy, we deal with multiple models. One way of categorizing them is the level at which they model the body and its interaction with radiation. We have models at the atomic and molecular levels (dose calculations), the molecular and cellular levels (radiobiology), the cellular and organ levels (TCP, NTCP) and the organism levels.

Another way to categorize our models is the degree to which they encapsulate cause-and-effect pathways (mechanistic models) versus empirical observations (probabilistic models). These categories are associated with the above-mentioned models of scale. The smaller the scale, the more likely the model is to represent specific mechanisms. The larger the scale, the more the models encapsulate clinical observations in the form of probability distributions. Clearly, the complexity of the environment being modeled influences the number of variables and the interactions between them.
Currently, we have a very accurate cause-and-effect model of the deposition of radiation at the atomic and molecular levels based on well-established theories of modern physics. Such exact knowledge fades quickly as we move to the biological effects of radiation absorption, even at the biomolecular and cellular level. Once basic ideas of free-radical-induced DNA strand breaks have been shown to incompletely describe what happens as we have learned more and more about the many molecular pathways of damage and repair. Now our models include very specific mechanisms, but also statistical distributions of measured parameters such as $\alpha$ and $\beta$. At the level of organ response, our mechanistic understanding is even less, with suboptimal concepts such as “parallel” and “serial” organs to guide us. Our models even incorporate their uncertain knowledge in their names, such as tumor control probability (TCP), and normal tissue complication probability (NTCP).

Obviously, if we are talking about models of our entire organism, we must rely on probabilistic models—assuming that such models even exist. Clinical exigencies force us to try to make predictions so that we can devise a course of treatment and counsel patients as to the decisions they must make. Recent developments have led to an emphasis on genomics as a source of knowledge that will allow us to develop “patient-specific treatments”. The amazing increase in our knowledge of the genetic components of disease and response gives us great hope that we are at the point where we can discard our probabilistic organism-level models for more exact mechanistic models.

In the following sections of this paper, we briefly describe why this may not occur as soon as we would like, and to give some examples of alternatives that we can adopt while we wait while our knowledge catches up to our needs.

2. Genomic models

A fundamental idea that has been the basis of most of theories of cancer development is that one rogue cell somehow escapes the normal bounds of cellular reproduction and growth and through unrestricted division leads to a cancer composed of identical genetic copies of the original cell. Such a simple view of the genetic structure of cancer is being rapidly replaced by a much more complex genetic system. Gerlinger et al studied renal cancers and found “intratumor heterogeneity and branched evolution revealed by multiregion sequencing” [1]. Vermaat et al looked at hepatic metastases of colorectal cancers and reported that the primary and metastatic tumors are genetically different which has many implications for the selection of patients for treatment targeted to a particular genotype [2]. For one, it calls into question the ability to determine the genetic range by means of any type of feasible biopsy. They hypothesize that such differences may lead to tumor adaptation and therapeutic failure through Darwinian selection. Clinicians are not surprised when multiple metastases respond completely differently to identical treatments, in opposition to the conventional wisdom that they are all identical. Another possibility is that different cellular milieus can have profound effects even on genetically identical tumor cells.

The idea that our current genetic knowledge is at a stage where we can develop such accurate models that genetically-targeted treatments are imminent is also proving less robust than some have predicted. For example, a paper by Nebert and Zhang found that most examples of pharmacogenetic traits reflect contributions from innumerable low-effect genes and explain only a small fraction of phenotype variation [3].

Similarly, in a study of monozygotic twins, Roberts et al performed whole genome sequencing and compared the incidence of disease of nine common cancers [4]. Their findings cast significant doubt that our current knowledge of genomics is going to be translated into patient-specific models of disease. Summarizing results from all nine cancers, they reported that only 0-40% of true positives would actually test positive; less than 10% of the general population would test positive; and there was only a 0-40% decrease in risk relative to the general population for those
testing negative.

Finally, as Liu points out, what is needed is not more lists of mutations, but rather, the ability to detect all relevant components of a system and to describe the complexity in a mathematical model [5]. He further discusses the need to reconstruct this complexity experimentally in order to learn more about their characteristics. Only that way can we make predictive models that explain possible responses of a cancer in terms of its gene networks.

We see that genetics certainly adds to our knowledge but we are still far from a mechanistic model at the organism level. We have no models that incorporates system responses, such as immune response. We also have no models that include environmental and stochastic variables. Even at the cellular level our models use surrogate variables that encapsulate much ignorance.

3. Probabilistic models

Randomized controlled clinical trials (RCT) are our best sources of data for probabilistic models. They also represent our best mechanistic models since it is usually the case that the design of an RCT is based on some prediction of cause and effect. The problem is that our knowledge is so incomplete that the model variables are very imprecise and many key variables and parameters are unknown. For example, three RTOG studies provide data for our models of the responses for the radiation treatment of brain metastases. There are really only three predictive variables (Karnofsky Performance Status, primary tumor status, and age), and the majority of patients actually fall into Class 2 because they don’t fall into the other two definitive categories.

Clinical trials are based on our best estimate of cause-effect relationships. The structure of the trial reflects the structure of the model. The results of the trial fill in the values of the model. Unfortunately, since there are so many unaccounted variables, the values are probability distributions over the variables of the trial. RCT’s are constrained to vary only one or two variables at a time, so that it takes many trials or other sources of information to fill in the entire model.

As stated before, one of the main uses of a model is to aid in decision making. In the case of medical treatment, these decisions are made under uncertainty. In these circumstances, we can consider a decison model as one that:

- predicts outcomes as a function of a given state and action;
- handles trade-offs between competing outcomes;
- provides some means of ranking outcomes.

Since this model must incorporate uncertainty, further qualities are:

- the use of probabilities as input;
- the ability to combine probabilities;
- the ability to perform inference.

Bayesian networks (BN) and their related structures, influence diagrams (ID), are one form of probabilistic model that are very appropriate for medical decision making [6, 7]. A BN is a directed acyclic graph that represents a joint probability table. It is constructed of variables connected by directed edges that imply causality. Each variable is characterized by more than one state (including continuous ones). There is a probability for an given state of any given variable that is conditioned on the states of each of its parent variables. These conditional probabilities are obtained from (a) randomized controlled trials, (b) medical research literature, (c) expert opinion, and (d) other models. Typically, the initial conditions of a medical situation are used to instantiate some of the variables while the probability calculus of the BN is used to calculate the probabilities of the other variables, usually some form of health outcome. If the model includes the evolution of the health states over time, it is often convenient to couple the
BN with a Markov model (MM) [8]. If the relative values of all the possible outcome states are measured by means of utilities, then the expectation value of the utilities can be calculated by using the probability distributions calculated by the BN, thereby creating an influence diagram [9]. In standard utility theory, the best decision is the one which leads to the maximum expected utility (MEU) [10].

The following sections provide three different influence diagrams that demonstrate different applications of these decision models in radiation therapy.

3.1. Decision model for selection of IMRT plans for prostate cancer
We have constructed an evolutionary multiobjective optimization algorithm to calculate a set of IMRT plans that embody the trade-offs between the competing objectives for prostate cancer [11, 12, 13]. In order to aid in the decision to select one from this set of deliverable plans, we constructed an influence diagram. Since one of the key objectives is to increase survival, and since prostate cancer can take years to develop, we included a Markov model to capture the time evolution [14, 15].

Figure 1 illustrates part of the BN (the nodes that represent normal tissue responses has been omitted for reasons of space). The probability of biochemical no evidence of disease (BNED) depends on the states of the three clinical variables and the delivered dose, as has been documented in a number of clinical trials. Other nodes and edges represent the decision variables used clinically and the conditional probabilities that they represent have been obtained from the literature.

Measuring patient-specific outcomes for the multiobjective health outcome states, e.g. \{tumor control\} + \{rectal condition\} + \{bladder condition\} + \{sexual function\}, is difficult. Using average measured utilities for populations of prostate cancer patients, we devised decision models for different classes of patients. The different classes were characterized by age and by their personal preference towards increasing life expectancy (tumor control) vs quality of life (complication-free). The combination of the influence diagram, Markov model and utilities allowed us to calculate the effect of different IMRT plans on the quality adjusted life expectancies, and to rank plans based on that metric.

3.2. Decision model for irradiation of occult lymph nodes in H&N cancer with and without FDG-PET
When designing treatment plans for head and neck cancer in the absence of definitive imaging or clinical evidence of positive lymph nodes, decisions need to be made whether or not to treat these...
“occult” nodes. Clinical experience has shown that in some percentage of cases, not treating node-negative regions will result in tumor recurrence. The probability of recurrence depends on the location and stage of the primary tumor, and these conditional probabilities are well-documented. Conventional CT imaging is often unable to visualize small tumors in the lymph nodes. It has been postulated that including FDG-PET imaging could improve the diagnosis of positive nodes. This influence diagram was constructed to model the effect on decision making of this additional information [16].

Possible actions in this model are whether or not to obtain PET imaging, what operating point in the receiver operator curve (ROC) designates a “positive”, and what dose to deliver. The ROC describes the accuracy of the information obtained from the PET images. The operating point sets the balance between false positives and false negatives. The selection of this point basically represents whether it is better to treat a tumor that isn’t there or to not treat a tumor that is.

The results of our decision modeling provided recommendations for different tumor sites and different stages, representing the varying probabilities of lymph node involvement. In some cases, the model recommends no use of PET since it would not change the clinical decision. This may seem counter-intuitive, but it correctly folds in the fact that there is some probability that the PET images are wrong. For example, late stage nasopharyngeal tumors result in lymph node metastases over 90% of the time. A negative PET image could represent a false negative and it is usually worth the risk of complications to treat that region. In other words, it is best to ignore the PET information in that setting. Other recommendations included setting of the operating point (usually towards high sensitivity) and the dose level.

An interesting result was the different way in which radiation oncologists view the same health state depending on whether it was the result of disease progression or the result of the radiation therapy. We found that it was necessary to separate out those two situations since clinicians often make conservative treatment decisions in order to avoid the possibility of complications. It is a personal decision as to whether reducing the chance of tumor control falls in the same category of “do no harm”.

3.3. Decision model in the design of a clinical trial of the efficacy of FLT-PET in re-irradiation of gliomas
Decisions regarding whether or not to re-treat a patient with a glioma are complicated by a number of factors, such as pseudo-progression and responses to other interventions. The question arose as to whether FLT-PET could be used to distinguish active regions of tumor
growth from other, non-malignant metabolic conditions. In other words, what information—what levels of sensitivity and specificity—would be necessary to change the clinical decisions and improve outcomes? A Bayesian network is an appropriate method of constructing a model to set bounds for the performance of PET in order to guide the design of a clinical trial.

Construction of the decision model is relatively straight-forward with respect to the variables and their states as well as in the connections between the variables. The difficulty lies in obtaining the probabilities. Much of the data in the literature has little in the way of a “gold standard” of truth with respect to the condition of the tumor. Thus it is hard to determine the conditional probabilities of the clinical examinations and the MR images.

There are established statistical methods for using a Bayesian approach to the design of clinical trials [17]. This differs from the method described here, although the statistical method can be part of an influence diagram. Basically, the difference has to do with the fact that the statistical approach can deal with a narrow range of relevant variables. The BN can incorporate all of the input variables and possible outcomes into the decision model, thereby providing a broader assessment of the potential influence of PET.

This work is on-going and, in the face of the challenges mentioned, the results are difficult to predict. However, as difficult as it is, it is easier than designing and carrying out trials without such guidance.

4. Conclusion

Hopes for an imminent mechanistic model of organism-level radiation responses based on improved understanding of genetics appears to be unrealistically optimistic. The situation is perhaps best summed up by Topol: “But these predictions will always remain estimates, and DNA’s power to predict illness will always be limited. The key to the future of individualized medicine is that such estimates and probabilities lead to actionable steps and prevention of disease.” [18] The need to make clinical decisions regarding of the our theoretic understanding compels us to make decision models as best we can. Currently, that means that our models are probabilistic. We have demonstrated that Bayesian networks and the related influence diagram are a productive means for constructing such models and we gave several examples of their use.

References

[1] Gerlinger M, Rowan A J, Horswell S, Larkin J and et al 2012 *N Engl J Med* 366 883–892
[2] Vermaat J S, Nijman I J, Koudijs M J, L G F and et al 2012 *Clin Cancer Res* 18 688–699
[3] Nebert D W and Zhang G 2012 *Science* 337 910
[4] Roberts N J, Vogelstein J T, Parmigiani G, Kinzler K W, vogelstein B and Velculescu V E 2012 *Sci Translat Med* 4 1–9
[5] Liu E T 2013 *Science* 339 1493
[6] Nease R F and Owens D K 1997 *Med Decis Making* 17 263–276
[7] Owens D K, Schachter R D and Nease R F 1997 *Med Decis Making* 17 241–262
[8] Beck J R and Pauker S G 1983 *Med Decis Making* 3 419–458
[9] Doctor J N and Miyamoto J M 2003 *J Mathematical Psychology* 47 557–567
[10] Von Neumann J and Morgenstern O 1944 *Theory of Games and Economic Behavior* (New York: Wiley)
[11] Holdsworth C, Kim M, Liao J J and Phillips M H 2010 *Med Phys* 37 4986–4997
[12] Holdsworth C, Stewart R, Kim M, Liao J J and Phillips M H 2011 *Med Phys* 38 2964–2974
[13] Holdsworth C H, Kim M, Liao J and Phillips M 2012 *Med Phys* 39 2261–2274
[14] Meyer J, Phillips M H, Cho P S, Kalet I and Doctor J N 2004 *Phys Med Biol* 49 1637–1653
[15] Smith W P, Doctor J, Meyer J, Kalet I J and Phillips M H 2009 *Artificial Intelligence in Medicine* 46 119–130
[16] Phillips M H, Smith W P, Parvathaneni U and Laramore G 2010 *Int J Radiat Oncol Biol Phys* 1089–1095
[17] Spiegelhalter D J, Abrams K R and Myles J P 2004 *Bayesian Approaches to Clinical Trials and Health-Care Evaluations* (John Wiley and Sons)
[18] Topol E J 2012 *Sci Translat Med* 4 135