6-1-2022

Why FLASH Radiotherapy Is Efficient: A Possible Explanation

Julio Urenda
The University of Texas at El Paso, jcurenda@utep.edu

Olga Kosheleva
The University of Texas at El Paso, olgak@utep.edu

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Nguyen Hoang Phuong
Thang Long University, nhphuong2008@gmail.com

Comments:

Technical Report: UTEP-CS-22-68

Why FLASH Radiotherapy Is Efficient: A Possible Explanation

Julio C. Urenda, Olga Kosheleva, Vladik Kreinovich, and Nguyen Hoang Phuong

Abstract Usually, a cancer radiotherapy session lasts between 10 to 20 minutes. Technically, it is possible to transmit the dose faster, but traditionally, medical doctors were reluctant to do it, since they were afraid of negative effects of such a speedy treatment. Recent experiments show, however, that these fears are unfounded; moreover, transmitting the whole radiation dose in a shorter time turns out to be more beneficial for the patients. In this paper, we provide a possible geometric explanation for this empirical phenomenon.

1 Formulation of the Problem

What is radiotherapy: a brief reminder. Radiotherapy is one of the most effective ways of fighting cancer, when radiation is applied to the tumor.

Traditional radiotherapy: in brief. Traditionally, each radiotherapy session lasts between 10 to 30 minutes.

From the purely engineering viewpoint, it is possible to apply the same overall dose of radiation faster. However, traditionally, medical doctors were reluctant

Julio C. Urenda and Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso
500 W. University, El Paso, Texas 79968, USA
e-mail: jcurenda@utep.edu, vladik@utep.edu

Olga Kosheleva
Department of Teacher Education, University of Texas at El Paso
500 W. University, El Paso, Texas 79968, USA
e-mail: olgak@utep.edu

Nguyen Hoang Phuong
Division Informatics, Math-Informatics Faculty, Thang Long University, Nghiem Xuan Yem Road
Hoang Mai District, Hanoi, Vietnam, e-mail: nhphuong2008@gmail.com

to bring this dose too fast, since they were afraid of undesirable side effects of radiation-per-hour.

FLASH radiotherapy. Recently, it turned out that not only these fears were unfounded: a very fast session, when the whole dose of radiation is sent in a few seconds, seem to work even better than the traditional radiotherapy.

Specifically, experiments on animals have shown that such *FLASH* radiotherapy actually decreases the undesirable effects on the surrounding healthy cells; see, e.g., [1, 2] and references therein. These results make FLASH radiotherapy a promising approach to cancer treatment.

Problem. How can we explain these unexpected experimental results?

What we do in this paper. In this paper, we provide a possible explanation for the success of FLASH radiotherapy.

2 Our Explanation

Let us formulate the problem in precise terms. The main objective of a radiotherapy session is to transmit the overall radiation dose D_0. How we do it can be characterized by a function $D(t)$ that describes how much radiation has been transmitted during this session by time t (starting from the beginning of this section).

We start at time $t = 0$, when the dose-so-far is 0: $D(0) = 0$. At each moment of time, we can only add more radiation, so this function is (non-strictly) increasing: if $t \le t'$ then $D(t) \le D(t')$. At the end of the session, we should have $D(t) = D_0$ – and this amount should remain the same until the next session.

Since the time until the next session – usually, several weeks – is much larger than the session's duration, it makes sense, when describing the current session, to ignore this future session – which is too far away from now – and to simply assume that the function $D(t)$ remains equal to D_0 for all non-negative value t.

In these terms, selecting an appropriate schedule means selecting a (non-strictly) increasing function $D(t)$ for which $D(0) = 0$ and

$$\lim_{t\to\infty} D(t) = D_0.$$

There are many such functions, which of them should we choose? Informally, we should select the best of these functions, the question is how we describe "the best" in precise terms.

How to describe "the best" in precise terms: general case. Usually, "the best" means that some objective function attains the largest (or the smallest) possible value. However, this is not the most general way of describing optimality. For example, if you have two different doses that have the same curing effect, it is reasonable to use this non-uniqueness to optimize something else: e.g., minimize the possibility of negative side effects. In this case, instead of the original single objective function

$f(a)$, we have a more complicated scheme, when an alternative a is better than an alternative b if either $f(a) > f(b)$, or if $f(a) = f(b)$ and $g(a) < g(b)$ for some other function $g(a)$.

If this more complex scheme still selects several alternatives, we can use this non-uniqueness to optimize something else, etc., until we reach the final optimality criterion in which we have only one optimal alternative.

The only thing we can say about such more general optimization settings is that we should be able, for any two alternatives a and b, to decide whether a is better than b (we will denote it by $a > b$), or b better than a ($b > a$), or a and b have the same quality (we will denote it by $a \sim b$). These relations $a > b$ and $a \sim b$ should satisfy natural consistency requirements: e.g., if a is better than b and b is better than c, then a should be better than c. Thus, we arrive at the following definition.

Definition 1. *Let A be a set. Its elements will be called* alternatives*.*

- *By an* optimality criterion*, we mean a pair of binary relations $\langle >, \sim \rangle$ that satisfy the following conditions for all a, b, and c:*
  - *if $a > b$ and $b > c$, then $a > c$;*
  - *if $a > b$ and $b \sim c$, then $a > c$;*
  - *if $a \sim b$ and $b > c$, then $a > c$;*
  - *if $a \sim b$ and $b \sim c$, then $a \sim c$;*
  - *if $a > b$, then we cannot have $a \sim b$.*
- *We say that an alternative a_{opt} is* optimal *with respect to the optimality criterion $\langle >, \sim \rangle$ if for every $a \in A$, we have either $a_{\text{opt}} > a$ or $a_{\text{opt}} \sim a$.*
- *We say that the optimality criterion is* final *if there exists exactly one alternative which is optimal with respect to this criterion.*

What are alternatives in our case. In our case, alternatives are different (non-strictly) increasing functions $D(t)$ for which $D(0) = 0$ and $D(t) \to D_0$ as $t \to \infty$.

Definition 2. *Let D_0 be a constant. By a D_0-*alternative*, we mean a (non-strictly) increasing function $D(t)$ for which $D(0) = 0$ and $D(t) \to D_0$ as $t \to \infty$.*

Natural invariance. There is no fixed unit of time relevant for this process, so it makes sense to require that the optimality criterion will not change if we use a different measuring unit to measure time.

If we know the dependence $D(t)$ in the original scale, how will this dependence look like in the new scale? If we replace the original measuring unit by a one which is λ times larger, then moment t in the new scale corresponds to moment $\lambda \cdot t$ in the original scale. For example, if we replace second with minutes – which are 60 times larger, then 2 minutes in the new scale is equivalent to $2 \cdot 60 = 120$ seconds.

In general, the value $D_{\text{new}}(t)$ corresponding to moment t in the new scale is thus equal to the value $D(\lambda \cdot t)$ when time is described in the original scale. Thus, $D_{\text{new}}(t) = D(\lambda \cdot t)$, and we arrive at the following definition.

Definition 3. *Let D_0 be a real number.*

- *For every $\lambda > 0$ and for every D_0-alternative $D(t)$, by a λ*-rescaling *$R_\lambda(D)$, we mean a D_0-alternative $D_{\text{new}}(t) \stackrel{\text{def}}{=} D(\lambda \cdot t)$.*
- *We say that the optimality criterion of the set of all D_0-alternatives is* scale-invariant *if for every $\lambda > 0$ and for every two D_0-alternatives a and b, we have the following:*
  - *if $a > b$, then $R_\lambda(a) > R_\lambda(b)$, and*
  - *if $a \sim b$, then $R_\lambda(a) \sim R_\lambda(b)$.*

Main result. Now, we are ready to formulate our main result.

Proposition. *Let D_0 be a real number, and let $(<, \sim)$ be a final scale-invariant optimality criterion on the set of all D_0-alternatives. Then, the optimal D_0-alternative has the form $D(t) = D_0$ for all $t > 0$.*

Discussion. This result explains the empirical fact that an instantaneous ("flash") radiotherapy indeed leads to the best medical results.

Proof.

1°. Let us first prove that for every final scale-invariant optimality criterion on the set of all D_0-alternatives, the optimal D_0-alternative D_{opt} is itself scale-invariant, i.e., $R_\lambda(D_{\text{opt}}) = D_{\text{opt}}$ for all $\lambda > 0$.

Indeed, by definition, the fact that D_{opt} is optimal means that for every D_0-alternative D, we have either $D_{\text{opt}} > D$ or $D_{\text{opt}} \sim D$. This is true for every D_0-alternative D, thus, this property holds for $R_{\lambda^{-1}}(D)$, i.e., we have either $D_{\text{opt}} > R_{\lambda^{-1}}(D)$ or $D_{\text{opt}} \sim R_{\lambda^{-1}}(D)$.

Since the optimality criterion is scale-invariant, we can conclude that either $R_\lambda(D_{\text{opt}}) > R_\lambda(R_{\lambda^{-1}}(D)) = D$ or $R_\lambda(D_{\text{opt}}) > R_\lambda(R_{\lambda^{-1}}(D)) = D$. This is true for all D_0-alternatives D. Thus, by definition of optimality, this means that the D_0-alternative $R_\lambda(D_{\text{opt}})$ is also optimal.

However, we assumed that our optimality criterion is final. This means that there is only one optimal D_0-alternative, and thus, $R_\lambda(D_{\text{opt}}) = D_{\text{opt}}$. The statement is proven.

2°. Let us now use the result from Part 1 of this proof to prove the Proposition, i.e., to prove that the optimal D_0-alternative has the desired flash form.

Indeed, the equality $R_\lambda(D_{\text{opt}}) = D_{\text{opt}}$ means that the values of these two functions coincide for all t. By definition of λ-rescaling, this means that for every t and every $\lambda > 0$, we have $D_{\text{opt}}(\lambda \cdot t) = D_{\text{opt}}(t)$. In particular, by taking $\lambda = s > 0$ and $t = 1$, we conclude that for every $s > 0$, we have $D_{\text{opt}}(s) = D_{\text{opt}}(1)$. Thus, the function $D_{\text{opt}}(s)$ attains the same constant value $D_{\text{opt}}(1)$ for all $s > 0$.

In particular, for $s \to \infty$, we have $D_{\text{opt}}(s) \to D_{\text{opt}}(1)$. By definition of a D_0-alternative, this limit must be equal to D_0. Thus, $D_{\text{opt}}(1) = D_0$ and therefore, for all $s > 0$, we have $D_{\text{opt}}(s) = D_0$.

The Proposition is proven.

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science), and HRD-1834620 and HRD-2034030 (CAHSI Includes), and by the AT&T Fellowship in Information Technology.

It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478, and by a grant from the Hungarian National Research, Development and Innovation Office (NRDI).

References

1. B. Lin, F. Gao, Y. Yang, D. Wu, Y. Zhang, G. Feng, T. Dai, and X. Du, "FLASH radiotherapy: history and future", *Frontiers in Oncology*, 2021, Vol. 11, Paper 644400.
2. P. Montay-Gruel, M. M. Acharya, O. Gonçalves Jorge, B. Petit, I. G. Petridis, P. Fuchs, R. Leavitt, K. Petersson, M. Gondré, J. Ollivier, R. Moeckli, F. Bochud, C. Bailat, J. Bourhis, J. F. Germond, C. L. Limoli, M. C. Vozenin, "Hypofractionated FLASH-RT as an effective treatment against Glioblastoma that reduces neurocognitive side effects in mice", *Clinical Cancer Research*, 2021, Vol. 27, No. 3, pp. 775–784.