Optimizing treatment combination for lymphoma using an optimization heuristic

Nicolas Houy, François Le Grand

Version of Record: https://www.sciencedirect.com/science/article/pii/S0025556418305741
Manuscript_64ef7e72fe7ad77c6b81018c365bee91

Optimizing treatment combination for lymphoma using an optimization heuristic*

Nicolas Houy† François Le Grand‡

July 8, 2019

Abstract

Background. The standard treatment for high-grade non-Hodgkin lymphoma involves the combination of chemotherapy and immunotherapy. We characterize in-silico the optimal combination protocol that maximizes the overall survival probability. We rely on a pharmacokinetics/pharmacodynamics (PK/PD) model that describes the joint evolution of tumor and effector cells, as well as the effects of both chemotherapy and immunotherapy. The toxicity is taken into account through ad-hoc constraints. We develop an optimization algorithm that belongs to the class of Monte-Carlo tree search algorithms. Our simulations rely on an in-silico population of heterogeneous patients differing with respect to their PK/PD parameters. The optimization objective consists in characterizing the combination protocol that maximizes the overall survival probability of the patient population under consideration.

Results. We compare using in-silico experiments our results to standard protocols and observe a gain in overall survival probabilities that vary from 4 to 9 percentage points. The gains increase with the complexity of the potential protocol. Gains are larger in presence of a higher number of injections or of an actual combination with immunotherapy.

Conclusions. In in-silico experiments, optimal protocols achieve significant gains over standard protocols when considering overall survival probabilities. Our optimization algorithm enables us to efficiently tackle this numerical problem with a large dimensionality. The in-vivo implications of our in-silico results remain to be explored.

Keywords: high-grade non-Hodgkin lymphoma, PK/PD model, protocol combination, Monte-Carlo tree search.

*We are very grateful to Michel Tod and Benoit You for their thorough comments.

†University of Lyon, Lyon, F-69007, France; CNRS, GATE Lyon Saint-Etienne, F-69130, France. Email: houy@gate.cnrs.fr.

‡emlyon business school, Écully, F-69130, France; ETH Zurich, Zurich, CH-8092, Switzerland. Email: legrand@em-lyon.com.

1 Introduction

Lymphoma is a cancer affecting the lymphatic system and in particular lymphocytes that tend to grow out of control. Since the lymphatic system involves many organs and covers the whole body, the lymphoma can be accompanied by a number of tumors throughout the body. There are two main types of lymphomas, Hodgkin and non-Hodgkin, where the former is characterized by the presence of giant cells derived from B-cells, the so-called Reed-Sternberg cells. Non-Hodgkin lymphomas are much more frequent than Hodgkin ones. The American Cancer Society [1] reports that the incidence rate of the Hodgkin lymphoma over the period 2010-2014 in the US is 2.7/100,000, while it reaches 19/100,000 for the non-Hodgkin lymphoma. Non-Hodgkin lymphomas (NHL, henceforth) covers under a unique denomination a wide variety of different types, that may start and affect different organs. NHL are also characterized by heterogeneous growth rates and different aggressiveness. On the one hand, low-grade lymphomas are said to be indolent and correspond to a slow development. On the other hand, high-grade lymphomas are said to be aggressive and experience a rapid development. All lymphomas are categorized in the Revised European-American Lymphoma Classification [11] – to which corresponds the World Health Organization classification [32].

The standard treatment for aggressive NHL relies on combination chemotherapy [9]. More precisely, the standard multi-drug combination – the so-called CHOP – involves cyclophosphamide (750 mg/m^2), doxorubicin (50 mg/m^2), vincristine (2 mg/m^2) on day 1, and prednisone (100 mg/m^2) given on days 1 to 5. The standard cycle amounts to 21 days and is denoted CHOP-21. In the 1980s, several trials involving chemotherapy combinations with up to 8 drugs have been conducted [6]. These different combinations turn out to offer a very limited improvement in survival rates while being both more toxic and more costly [13, 30]. This was confirmed in a randomized phase III trial, where different combinations were tested [10]. The German High-Grade Non-Hodgkin's Lymphoma Study Group has analyzed several variations around the standard CHOP-21. The main variations are: (i) the intensification of the cycle, from 21 to 14 days, (ii) the introduction of an additional drug through the administration of etoposide, a potent cytotoxic agent, and (iii) the intensification of drug doses in chemotherapy. Findings conclude only to partial success. Either intensifying the cycle or introducing etoposide in a 21-day cycle (corresponding to the so-called CHOEP-21) yields an increase in event-free survival rates [22–24]. Therefore, a moderate intensification in cycles or in doses leads to better outcomes. However, increasing further doses or cycles deteriorate survival rates. For instance, the cycle intensification together with etoposide introduction – CHOEP-14 – is more toxic and leads to lower

survival rates [23]. Similarly, doubling the drug doses with etoposide and a 21-day cycle does not provide any clinical benefit for young patients [26].

Furthermore, a paradigm shift in the treatment of NHL occurred with the combination of standard chemotherapy (CHOP and CHOEP) with immunotherapy through the injection of rituximab, which is a monoclonal antibody [5, 25]. Rituximab contributed to a 50% reduction in patient mortality, which makes it one of the largest success in NHL treatment for the last years [17]. More precisely, rituximab is a type-I anti-CD20 antibody that binds to CD20 receptors on the surface of B-cells [27]. This triggers complement-dependent cytotoxicity, which eliminates B-cells by direct lysis [20]. This cytotoxic effect is completed by another mechanism, the antibody dependent cellular cytotoxicity. This mechanism makes activated B-cells visible to immunosuppression by effector cells bearing Fc receptors [33]. Furthermore, if rituximab has only a minor direct impact on apoptosis, it makes cancer cells more sensitive to chemotherapy [2, 4]. As for the sole chemotherapy treatment, the question of the cycle and dose intensities has also to be investigated in case of the combination treatment. A cycle of 14 days for the combination of rituximab with standard CHOP (R-CHOP, henceforth) is shown to bring no improvement over a cycle of 21 days [7]. A similar absence of effects has been shown for the intensification of the doses in a R-CHOP21 regimen [8]. The question of designing an optimal combination protocol remains open [18, 21]. Obviously, given the very number of possibilities to be tested, in-silico trials and computational simulations can be helpful in providing protocol guidelines.

In this paper, we develop an optimization heuristic algorithm to characterize the in-silico protocol – with and without rituximab administration – that maximizes the overall survival rates in a population of heterogeneous fictive patients, subject to toxicity constraints. Our optimization relies on in-silico simulations that are based on the paper by Roesch *et al.* [29]. These authors provide a model of the development of high-grade NHL that takes into account the combined cytotoxic effects of chemotherapy and immunotherapy. We rely on population variability data from [16, 28, 29]. Our optimization objective consists in characterizing the optimal injection schedule that maximizes in-silico the number of overall survivals in a population of heterogeneous patients, subject to some ad-hoc toxicity constraints. The constraints we impose bear on the minimal time distance between two injections – for chemotherapy or immunotherapy –, as well as the total number of injections. The objective of these constraints aims at proxying drug toxicity. In particular, we make no use of an explicit physiological modelling of toxicity. We distinguish two cases, depending on whether immunotherapy is introduced or not. The patients differ according to their immunotherapy pharmacokinetics, as well as to their pharmacodynam-

ics. We compare the outcomes of our optimal protocol to those of the standard CHOP protocols. Our results unambiguously deliver higher survival rates for in-silico simulations. Compared to standard protocols, the progression in survival rates of our fictive population amounts to 5 percentage points (pp) in absence of immunotherapy and even to 9 pp when immunotherapy is introduced. In the same fictive population, the survival rates increase from 31.0% to 36.5% in absence of immunotherapy, and from 47.5% to 56.2% in presence of immunotherapy. Therefore, optimizing the joint schedule of dose injections enables to significantly improve the outcome of treatment, at least in numerical simulations. The question of the in-vivo implications of these in-silico results remains open. It is obviously subject to the usual limitations of a study based on a theoretical model and on prospective data.

These results also illustrate the potential gains of introducing such an optimization heuristic algorithm in computational oncology and in particular for characterizing in-silico the optimal combination protocols involving both chemotherapy and immunotherapy. The algorithm we rely on belongs to the family of Monte-Carlo Tree Search (MCTS, henceforth) algorithms. MCTS methods are surveyed in [3]. This family of algorithms is well-known for their application in a Go playing software. The famous AlphaGo has been very successful in defeating several Go champions in 2016-2017. The applications of MCTS methods to Go software can be found in [31]. We provide a detailed explanation of our algorithm in Section 3.2. To the best of our knowledge, MCTS algorithms have not been applied to combinations of chemotherapy and immunotherapy. However, such algorithms have already been used successfully to characterize optimal protocols, both in chemotherapy [12] and in immunotherapy [19] (but not with a treatment combination). One of the advantages of MCTS algorithms is that they enable to handle a full-fledge PK/PD model, without any need to simplify it. They take advantage of the sequential nature of treatments in oncology and also allow for a very high degree of flexibility for real-life implementation.

2 Materials and methods

2.1 PKPD model

We present here the model which our optimization exercise relies on. This model has been proposed by Roesch et al. [29]. It builds on a model by Kuznetsov et al. [14] for chemotherapy and initially intended for modelling leukemia in mice. The model relies on two main ordinary differential equations driving the evolution of the number of tumor cells and of effector cells. The latter cells can be thought of as corresponding to CD8+ cytotoxic T-cells. In absence of chemotherapy and of effector cells, the tumor follows

an exponential uncontrolled growth. Effector cells interact with tumor cells and their interaction is proportional to the surface of the tumor. The interaction of both cell types leads to the killing of tumor cells. The production of effector cells is fostered by tumor cells through immunogenicity. Effector cells die naturally and after interaction with tumor cells. In absence of treatment, the evolution of the cancer depends on which type of cells dominate the other. There is cure if tumor cells deplete and is overcome by effector cells. On the opposite, there is no cure if tumor cells dominate effector cells. Chemotherapy negatively affects the tumor growth, but also effector cells. Chemotherapy is therefore a double-edge sword and can have the paradoxical effect to foster tumor cells by killing too many effector cells. Finally, rituximab also has a twofold effect. First, it has a direct cytotoxic effect on tumor cells only – not on effector cells. Second, it boosts the immunogenicity effect and fosters the growth of effector cells.

For the calibration of the model, we rely on the parameters provided in [28, 29], as well as [16] for the pharmacokinetics of immunotherapy. Of note, these parameter values include population variability. A detailed presentation of the model, with the parameter values that we use, can be found in Appendix A.

2.2 Simulations

We simulate the model evolution for a population of 10,000 heterogeneous patients. The model parameters of these patients are randomly drawn in the distribution of parameters that reflects the population variability. For each patient, we simulate the evolution of the tumor cells and effector cells until the number of tumor cells reaches a diagnosis threshold which is itself a random variable (see Appendix A for further details). We consider this state – when the number of tumor cells reaches the diagnosis threshold – as the initial state. Notice that because some parameters are patient-dependent, the initial state is also patient-dependent. We plot in Figure 1 the initial states (i.e., the numbers of tumor and effector cells at diagnosis) in the population of patients. Note that for some parameter values (in the dynamic system but also the diagnosis threshold itself), cancer is never diagnosed. These individuals are discarded and therefore do not appear in Figure 1. This explains the apparently complex distribution of initial values.

Then, the simulation is assumed to start at diagnosis date and ends 730 days – two years – after the end of the treatment. We make this choice to avoid that the length of the treatment influences our results. Indeed, in our simulations, increasing further the recovery period does not change our results since the dynamic system has already converged (or diverged).

A patient will be considered to be cured if the number of tumor cells at the final date

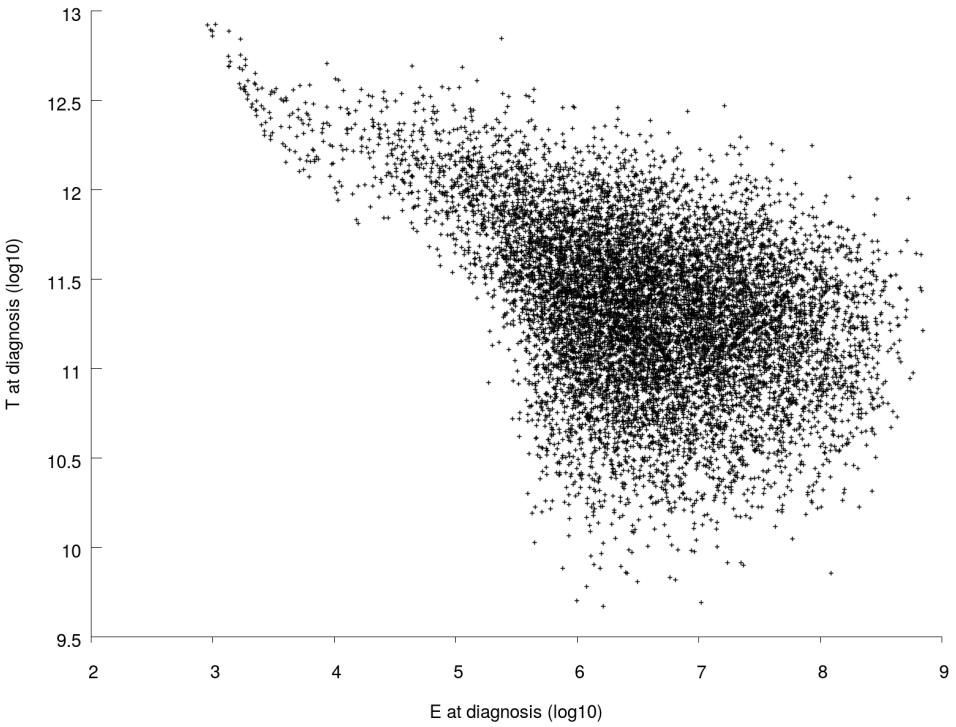


Figure 1: Distribution of the number of tumor cells and of effector cells in the initial population for patients whose cancer has been diagnosed.

is below one cell. Oppositely, a patient will be said to have a progressive disease if the number of effector cells ends up being below one. This case also corresponds to a very large number of tumor cells, typically above 10^{13} cells. Besides these two polar cases, there is also the possibility that the number of tumor and effector cells both remain above 1 two years after the end of the treatment. This last case corresponds to a cancer where the cytotoxic effect of effector cells offsets the growth of the tumor for at least two years after the end of the treatment. In our simulations, this last case is not very frequent and typically represents less than 0.1% of patients.

To assess the efficacy of any protocol, we proceed as follows. We assume that the protocol under consideration is administered to the whole population and we then compute the proportion of patients, who end up with a cured cancer at the end of the simulation horizon, after two-years. We will henceforth refer to this share as the two-year overall survival probability. Of note, the patients with a stable cancer for two years will not be considered as having survived. This choice is a bit conservative but has no major influence on results since the proportion of such cancers is very low.

All our simulations are implemented in C++.

3 Results

We compute the outcomes of two families of protocols. In the first one, we consider protocols based on a cycle. In the second family, we characterize using MCTS algorithms the non-cyclic protocols achieving the highest two-year overall survival probability.

3.1 Cycle protocols

We focus here on the protocols based on a cycle. For the sake of simplicity, for chemotherapy we denote xCHOPy the protocol with x the number of cycles and y the cycle length in days. Similarly for immunotherapy, we consider protocols of the form xRy, where x and y have the same signification as for chemotherapy (number of cycles and cycles length). A protocol that will combine chemotherapy and immunotherapy will therefore combine both protocols CHOP and R. As an illustration, we report in Table 1 the outcomes for some standard cycle protocols. For the survival probability, we report the observed value for our in-silico population of 10,000 individuals as well as the implied 95% confidence interval (computed as a Wilson score interval). For instance, the protocol with 12 biweekly CHOP cycles and 14 21-day long ritumixab cycles (12CHOP14-14R21) leads to survival for 47.52% of our population patients.

Chemotherapy	Immunotherapy	Two-year overall survival probability (%)
6CHOP14	None	28.60 [27.72−29.49]
6CHOP14	8R21	42.53 [41.56−43.50]
6CHOP14	14R21	42.83 [41.86−43.80]
12CHOP14	None	30.96 [30.06−31.87]
12CHOP14	8R21	44.80 [43.83−45.78]
12CHOP14	14R21	47.52 [46.54−48.50]

Table 1: We consider various protocols combining chemotherapy and immunotherapy. For the two-year overall survival probability, we report the observed value and below, between square brackets, the implied 95% confidence interval.

One of the main lessons of Table 1 is that including immunotherapy in a standard chemotherapy protocol drastically improves treatment outcomes. This is particularly striking when comparing the rows in absence of immunotherapy (rows 1 or 4) to their respec-

tive counterparts in presence of immunotherapy (rows 2 and 3 or 5 and 6, respectively). The in-silico model reflects the significant benefits from combining immunotherapy with chemotherapy observed in clinical trials [17].

We focus here on protocols with chemotherapy only. We compute for any cycle length the interval between injections that leads to the highest survival probability. More formally, for any $x \in [\![1, 12]\!]$, we determine $y^* \in [\![7, 35]\!]$, such that $x\text{CHOP}y^*$ yields a higher survival probability than any other protocol $x\text{CHOP}y$ ($y \in [\![7, 35]\!]$).[1] We plot in Figure 2 the survival probabilities associated to these optimal protocols with cycles. Exact values (as well as the cycle length maximizing efficacy) can be found in Appendix B.1. Two lessons can be drawn from Figure 2. First, a moderate increase in the number of cycles significantly improves the protocol outcome. Increasing the number of cycles from 1 to 8 increases the survival probability from 4.4% to 32.5%. Second, increasing the number of cycles further does not yield any further outcome improvement. We can even observe a slight degradation of survival probabilities. For instance, moving from 8 to 12 cycles slightly diminishes the two-year overall survival probability from 32.5% to 31.3%. These in-silico findings are consistent with clinical trials, showing that if cycle intensification has positive effects for moderate total cycle numbers, the intensification turns out to generate negative effects for large cycle cycle numbers [23].

Given the very high number of possible combinations, it is not possible to replicate the graph in Figure 2 for the combination of immunotherapy and chemotherapy. Yet, we will be able to use these results as a benchmark for comparison with non-cyclic protocols.

3.2 Unconstrained optimization

We now focus on unconstrained protocols. Using a MCTS algorithm, we determine the non-cyclic protocols that maximize the two-year overall survival probability for a population with the same parameters distributions as ours. Then, the population on which we test our optimized protocol has the exact same characteristics as the one in Sections 2.2 and 3.1. These optimal protocols are parametrized by the number of chemotherapy injections and the number of immunotherapy injections. We formally denote such an optimal protocol $OC_{x,y}$ where OC stands for optimal combination, x the number of chemotherapy injections and y the number of immunotherapy injections. We furthermore impose two constraints in our optimization. First, the time interval between two chemotherapy injections should not be smaller than 7 days. Second, we impose the same time constraint for immunotherapy injections. This limit of 7 days stems from the fact that it corresponds to the minimal cycle length in [29]. The objective of these constraints is to mimic toxicity constraints.

[1] The set $[\![n, m]\!]$ gathers all integers from n to $m \geq n$.

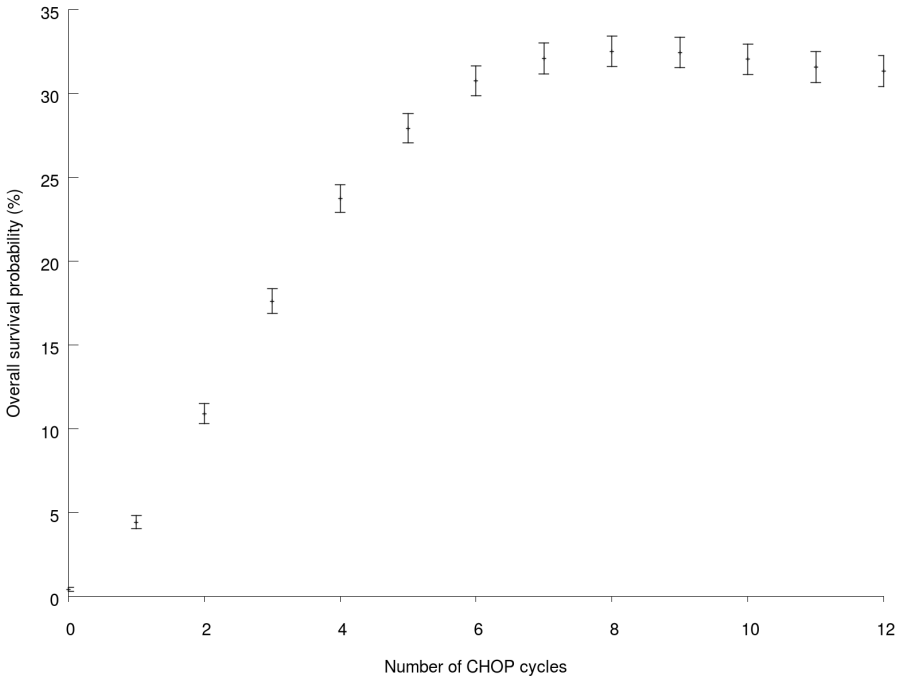


Figure 2: For each number of CHOP cycles (without immunotherapy), we determine the cycle length that leads to the highest survival probability. The point corresponds to the average and the two bars delimit the 95% confidence interval.

Indeed, in absence of physiological model of toxicity, we need to rely on proxies. This is a limitation of our exercise and should be kept in mind when interpreting the results and thinking about their in-vivo implications.

We report some results in Table 2 that mirror those of Table 1. Compared to Table 1 , it is straightforward to observe that optimal protocols lead to higher survival probabilities than standard combination protocols. Gains in two-year overall survival probabilities vary from 5 to 9 percentage points. Unsurprisingly, the gain of optimization is rising with possible complexity of the protocol. This complexity can be generated either by the combination of immunotherapy with chemotherapy or by the higher number of injections. Note that we discuss below the dosage of the different optimal protocols.

It is noteworthy that the dimensionality of the optimization exercise is very high. In particular, relying on a standard optimization technique such as dynamic programming is not an option here. Our algorithm enables to overcome this difficulty and to tackle the dimensionality issue. How does it work exactly? A detailed account of our algorithm in a simpler set-up featuring chemotherapy only (but no immunotherapy) is provided in [12]. We can sum up the functioning of our algorithm in the current set-up as follows. We start with a heterogeneous in-silico patient population, as described in Section 2.2. We consider

Optimal protocol	Two-year overall survival probability (%)
$OC_{6,0}$	32.71 [31.80−33.64]
$OC_{6,8}$	48.13 [47.15−49.11]
$OC_{6,14}$	49.69 [48.71−50.67]
$OC_{12,0}$	36.46 [35.52−37.41]
$OC_{12,8}$	54.06 [53.08−55.04]
$OC_{12,14}$	56.20 [55.23−57.17]

Table 2: For some optimal protocols, we report for the two-year overall survival probability the observed value in the patient population and below, between square brackets, the implied 95% confidence interval. The latter is again computed as a Wilson score interval.

also as given a set of schedule constraints (mimicking toxicity constraints: no less than 7 days between two consecutive administrations of either chemotherapy or immunotherapy), as well as constraints on the number of maximal administrations (for $OC_{6,8}$, no more than 6 chemotherapy and 8 immunotherapy administrations). The initial date is normalized at date 0.

At day 0, and as at any day, a treatment is possible and the algorithm must then decide upon the treatment on that day. In this PK/PD model, doses are fixed, so the algorithm needs to decide between two alternatives: injection or no injection for chemo- and immunotherapy, which adds up to 4 different possible combinations. Some combinations can be ruled out because of schedule constraints or because of the maximal bound on the number of administrations. Here obviously, at day 0, since no administration has taken place yet, none of these constraints is likely to be binding. The algorithm then computes the best combination among the four. To do so, each of the four combination is associated to a fictive population, which is an exact copy of the current population for which a decision needs to be made. The algorithm then simulates the evolution of the the number of effector and tumor cells for the four fictive populations, assuming that each population is administered a default continuation treatment policy (see below for further details about this default policy). The recommended combination is then the one corresponding to the population with the best outcome after simulation. The best outcome is assumed to be the one with the smallest number of tumor cells. The actual patient population is then administered the recommended combination. We then compute the evolution of the

number of tumor and effector cells in the population using the PK/PD model until the next day, which is day 1. The same procedure as in day 0 then applies. The algorithm checks the validity of schedule constraints and selects the best treatment combination. This is repeated every day until the last treatment day (day 180) and evolution of the population is then simulated – with no treatment administration – until day 730. The last simulation window guarantees that the dynamic system has converged.

The tricky part of the algorithm is the choice of default continuation treatment policy when simulating the four fictive patient populations. In this case, we choose the default policy associated to the optimal algorithm under consideration. For instance, for the computation of $OC_{6,8}$, the default policy will be the cycle protocol 6CHOP14–8R21.

Contrary to our previous studies [12, 19], our current exercise involves a two-dimensional choice (chemotherapy *and* immunotherapy), which corresponds to treatment combination. Furthermore, the dynamic system of the current PK/PD model is highly non-linear, as can be seen from Figure 4 below. This paper therefore shows that our optimization method can therefore handle a complex problem, with potential real-life applications.

4 Discussion

We now compare further into detail two protocols, 6CHOP14-8R21 and $OC_{6,8}$, with 6 injections of chemotherapy and 8 injections of immunotherapy each. In Table 3, we show the share of individuals in our in-silico population that are cured for each protocol, 6CHOP14-8R21 and $OC_{6,8}$.[2] As we have already shown, the optimization increases the two-year overall survival probability by almost 6 percentage points, since it moves from 42.53% to 48.13%. This net gain does not hide a large share of patients that suffer from the optimization. Only 0.67% of patients that were cured with the standard protocol are not cured with the optimized one. So, the net gain is almost equal to the share of patients cured by the optimized protocol that are not cured with the standard one. This last share amounts to 6.27%. In our in-silico simulations, the optimization therefore benefits to almost every patient.

We now turn to the dynamic results. In Figure 3, we plot the evolution of the overall survival probability from the date of diagnosis (and hence treatment) until 2 years after the last injection shot – that is attached to $OC_{6,8}$. The black line corresponds to 6CHOP14-8R21 and the grey line to $OC_{6,8}$. The graph makes it clear that the overall survival probability for both protocols has converged and that the last date value we consider does not influence results. We first observe that the two-year overall survival probability is

[2]Notice that both treatments are implemented for exactly the same in-silico populations.

	Cured 6CHOP14-8R21	Not cured 6CHOP14-8R21	Total
Cured $OC_{6,8}$	41.86%	6.27%	48.13%
Not cured $OC_{6,8}$	0.67%	51.20%	43.87%
Total	42.53%	57.47%	100%

Table 3: Proportions of patients that are cured and not cured with 6CHOP14-8R21 and $OC_{6,8}$ protocols.

higher for $OC_{6,8}$ than for 6CHOP14-8R21, which confirms the results of Table 3. The comparison between the two curves can be split in three parts. First, we observe that during the first 4 months of the treatment, the standard and optimized protocols perform very similarly. Second, between day 120 and day 500, the gap between the outcomes of both protocols progressively widen. The over-performance of optimized over standard protocol progressively reveals during that period. Third, after day 500, there is no further change in overall survival probabilities and the gap between both protocols remains unchanged to its value of 6 points approximately. Consequently, the gains of optimization in in-silico simulations only progressively appear and do so quite late in the treatment. Of note, Figure 3 shows that the model does not feature any late relapse, while they have been documented in the literature [10, 15].

We also report administration schedules in Table 4. Regarding chemotherapy, the treatment is very dense for the first three injections of $OC_{6,8}$, while the last three injections occur with a greater time interval than with 6CHOP14-8R21. The chemotherapy treatments of both protocols end up at close dates. We observe a similar pattern for immunotherapy administration. The first four injections occur on an intense schedule (within 22 days), while the four last ones are realized with much larger intervals. The last immunotherapy administrations for both the standard and optimal protocols are very close in time. We report in Section B.2 of the Appendix the administration schedules and the evolution of overall survival probabilities for other optimal protocols ($OC_{6,0}$, $OC_{6,14}$, $OC_{12,0}$, $OC_{12,8}$, and $OC_{12,14}$).

We now study the evolution of the cancer through the number of tumor and effector cells. We represent such an evolution for two specific patients in Figure 4. Panel 4a corresponds to a patient who is cured with $OC_{6,8}$ but not with 6CHOP14-8R21, and vice-versa for the patient of Panel 4b. These graphs show the highly non-linear dynamics of the cancer evolution. To better understand why, note that the trajectory for each protocol is

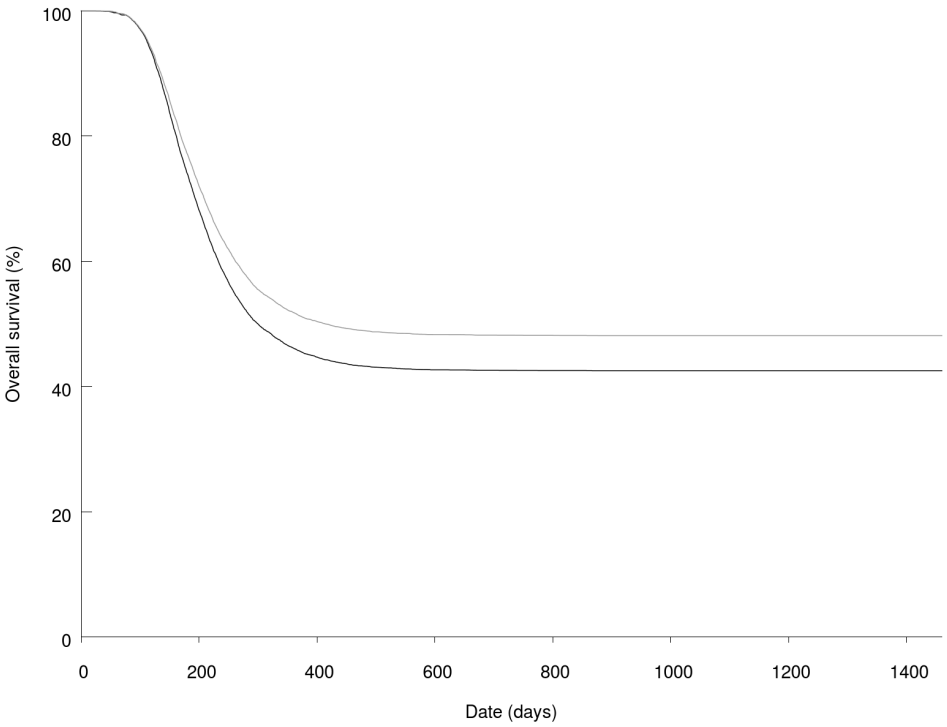


Figure 3: The two curves represent the evolution of the overall survival probabilities. The dark line corresponds to 6CHOP14-8R21 and the light color to $OC_{6,8}$.

made of two parts. The first part, which is a broken line, consists of the treatment itself, where each kink corresponds to an injection. The second part, which is a smooth line, corresponds to the evolution of the system after the treatment. On Panel 4a, the patient receiving $OC_{6,8}$ ends up with a smaller number of tumor cells, and a larger number of effector cells than the standard treatment. Overall, the optimized protocol yields tumor elimination, while the standard protocol does not. On Panel 4b, the situation is partly reversed: the optimized treatment ends up with slightly more effector cells, but less tumor cells. This time, only the standard protocol eliminates the tumor. So, given the highly non-linear dynamics of the cancer tumor, focusing only at one indicator at the end of the treatment can be misleading. A 'low' number of tumor cells does not necessarily imply tumor elimination eventually. The strength of our optimization algorithm is to be able to overcome these difficulties and to deliver an efficient protocol.

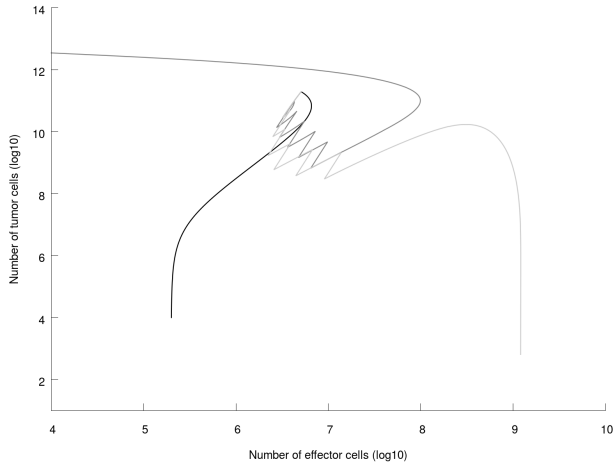
The variations are computed as follows. At each date of administration of chemo (resp. immuno), with a given shifting probability, we add a 1-day or 3-day shift for the current administration and the remaining of the chimo (resp. immuno) schedule. Shifting all subsequent administration dates enables to preserve the schedule constraints. Our results for the two-year overall survival probability are gathered in Table 5.

Day	6CHOP14-8R21	$OC_{6,8}$
1	chemo + immuno	chemo + immuno
8		chemo + immuno
15	chemo	chemo + immuno
22	immuno	immuno
26		chemo
29	chemo	
38		immuno
43	immuno+chemo	
44		chemo
57	chemo	
58		immuno
64	immuno	
65		chemo
71	chemo	
85	immuno	
106	immuno	
127	immuno	
123		immuno
148	immuno	
158		immuno

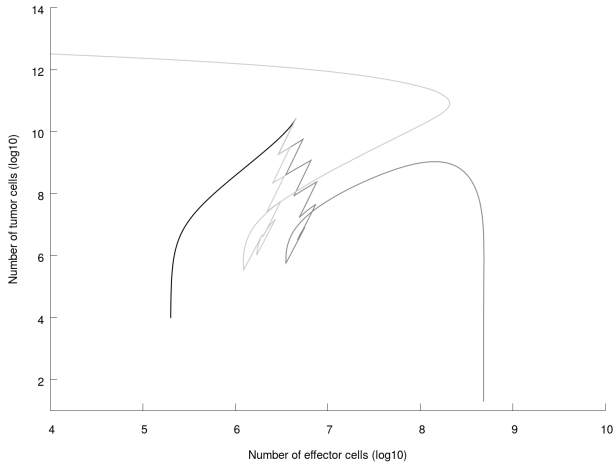
Table 4: Administration schedules of the protocols 6CHOP14-8R21 and $OC_{6,8}$.

Finally, as a robustness check, we study to which extent our results are sensitive to the exact timing of administration. Such a robustness check enables us to measure how real-life contingencies (missed administration because of a patient absence for instance) could affect protocol outcomes. We consider the optimal protocol $OC_{6,8}$ as our benchmark. We then perform a stochastic sensitivity analysis, in which we implement a random variation of $OC_{6,8}$ for each of the 10,000 patients of our in-silico population. A random variation of $OC_{6,8}$ is computed as follows. At each date of a chemotherapy (resp. immunotherapy) administration, we shift by 1 day or 3 days, with a given probability, the current administration date and all remaining chemotherapy (resp. immunotherapy) administration dates. Shifting all subsequent administration dates enables us to preserve the schedule constraints. Note that we exclude any variation that is identical to $OC_{6,8}$. Our results for the two-year overall survival probability are gathered in Table 5.

For sake of simplicity, the first row of Table 5 corresponds to a zero shift probability and therefore to $OC_{6,8}$. As can be observed, shifting administration days deteriorates



(a) Example of patient who cures with $OC_{6,8}$ but not with 6CHOP14-8R21.



(b) Example of patient who cures with 6CHOP14-8R21 but not with $OC_{6,8}$.

Figure 4: Dynamic evolution in the space of number of tumor and effector cells for two different patients. For both graphs, the dark line corresponds to the evolution before the diagnosis (with 10^4 tumor cells at initial time), the light line to the evolution for the 6CHOP14-8R21 protocol and the lightest line to the evolution for the $OC_{6,8}$ protocol.

the protocol outcome and the effect is larger when the probability is high and the time shift long. This conclusion is rather intuitive and consistent with the idea that $OC_{6,8}$ is optimal. However, the impact is relatively modest, at least for the 1-day time shift. For the three probabilities, the decrease in two-year overall survival probability is smaller than 1 point and is not significant. For the 3-day time shift, the impact remains small and non significant for a shifting probability of 10%. However, for the probabilities of 20% and 50%, this does not hold any more. In particular, for the 50% probability, the impact amounts to 3 points approximately, and is significant.

In in-silico trials, MCTS algorithms have proved to be helpful in characterizing efficient protocols for a treatment combination of chemotherapy and immunotherapy. The gain is sizable and amounts to almost 6 percentage points in two-year overall survival probabilities for a combinations with 6 chemotherapy injections and 8 immunotherapy ones – as in the non-optimized standard 6CHOP14-8R21. Gains can be larger for more complex combinations of chemo- and immunotherapy. However, when interpreting our results, two caveats should be kept in mind. First, in our in-silico optimization, toxicity is limited by ad-hoc constraints on time intervals between injections and not by an actual physiological toxicity model. Second, our results have been validated in-silico trials only. The underlying model is indeed likely to feature a number of simplifications that could affect outcomes (such as the absence of late relapses in the model, as discussed for Figure 3). As for any numerical

Shifting probability	Two-year overall survival probabilities	
	1-day shift	3-day shift
0	48.13 [47.15−49.11] ($= OC_{6,8}$)	
10%	47.86 [46.88−48.84]	47.42 [46.44−48.4]
20%	47.83 [46.85−48.81]	46.97 [45.99−47.95]
50%	47.35 [46.37−48.33]	45.23 [44.24−46.21]

Table 5: Robustness analysis of our results. Two-year overall survival probabilities (with 90% confidence intervals between brackets) when administration dates of $OC_{6,8}$ have been shifted by 1 or 3 days with a probability of 10%, 20% or 50%.

analysis, further work is needed to understand to which extent our conclusions would hold in-vivo.

5 Compliance with ethical standards

Conflict of interest

Both authors declare that they have no conflict of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

[1] The American Cancer Society. `https://cancerstatisticscenter.cancer.org/`. Accessed: 2018-07-11.

[2] S. Alas, C. Emmanouilides, and B. Bonavida. Inhibition of interleukin 10 by rituximab results in down-regulation of bcl-2 and sensitization of B-cell non-Hodgkin's lymphoma to apoptosis. *Clin. Cancer Res.*, 7(3):709–723, Mar 2001.

[3] C. Browne, E. Powley, D. Whitehouse, S. Lucas, P. I. Cowling, S. Tavener, D. Perez, S. Samothrakis, and S. Colton. A survey of monte carlo tree search methods. *IEEE Trans. Comput. Intellig. and AI in Games*, 4(1):1–43, 2012.

[4] K. U. Chow, W. D. Sommerlad, S. Boehrer, B. Schneider, G. Seipelt, M. J. Rummel, D. Hoelzer, P. S. Mitrou, and E. Weidmann. Anti-CD20 antibody (IDEC-C2B8, rituximab) enhances efficacy of cytotoxic drugs on neoplastic lymphocytes in vitro: role of cytokines, complement, and caspases. *Haematologica*, 87(1):33–43, Jan 2002.

[5] B. Coiffier, E. Lepage, J. Briere, R. Herbrecht, H. Tilly, R. Bouabdallah, P. Morel, E. Van Den Neste, G. Salles, P. Gaulard, F. Reyes, P. Lederlin, and C. Gisselbrecht. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N. Engl. J. Med.*, 346(4):235–242, Jan 2002.

[6] M. Coleman. Chemotherapy for large-cell lymphoma: optimism and caution. *Ann. Intern. Med.*, 103(1):140–142, Jul 1985.

[7] D. Cunningham, E. A. Hawkes, A. Jack, W. Qian, P. Smith, P. Mouncey, C. Pocock, K. M. Ardeshna, J. A. Radford, A. McMillan, J. Davies, D. Turner, A. Kruger, P. Johnson, J. Gambell, and D. Linch. Rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisolone in patients with newly diagnosed diffuse large B-cell non-Hodgkin lymphoma: a phase 3 comparison of dose intensification with 14-day versus 21-day cycles. *Lancet*, 381(9880):1817–1826, May 2013.

[8] R. Delarue, H. Tilly, N. Mounier, T. Petrella, G. Salles, C. Thieblemont, S. Bologna, H. Ghesquieres, M. Hacini, C. Fruchart, L. Ysebaert, C. Ferme, O. Casasnovas, A. Van Hoof, A. Thyss, A. Delmer, O. Fitoussi, T. J. Molina, C. Haioun, and A. Bosly. Dose-dense rituximab-CHOP compared with standard rituximab-CHOP in elderly patients with diffuse large B-cell lymphoma (the LNH03-6B study): a randomised phase 3 trial. *Lancet Oncol.*, 14(6):525–533, May 2013.

[9] V. T. DeVita, G. P. Canellos, B. Chabner, P. Schein, S. P. Hubbard, and R. C. Young. Advanced diffuse histiocytic lymphoma, a potentially curable disease. *Lancet*, 1(7901):248–250, Feb 1975.

[10] R. I. Fisher, E. R. Gaynor, S. Dahlberg, M. M. Oken, T. M. Grogan, E. M. Mize, J. H. Glick, C. A. Coltman, and T. P. Miller. Comparison of a standard regimen (CHOP) with three intensive chemotherapy regimens for advanced non-Hodgkin's lymphoma. *N. Engl. J. Med.*, 328(14):1002–1006, Apr 1993.

[11] N. L. Harris, E. S. Jaffe, H. Stein, P. M. Banks, J. K. Chan, M. L. Cleary, G. Delsol, C. De Wolf-Peeters, B. Falini, and K. C. Gatter. A revised European-American classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group. *Blood*, 84(5):1361–1392, Sep 1994.

[12] N. Houy and F. Le Grand. Optimal dynamic regimens with artificial intelligence: The case of temozolomide. *Plos One*, 13(6), 2018.

[13] P. Klimo and J. M. Connors. MACOP-B chemotherapy for the treatment of diffuse large-cell lymphoma. *Ann. Intern. Med.*, 102(5):596–602, May 1985.

[14] V. A. Kuznetsov, I. A. Makalkin, M. A. Taylor, and A. S. Perelson. Nonlinear dynamics of immunogenic tumors: parameter estimation and global bifurcation analysis. *Bull. Math. Biol.*, 56(2):295–321, Mar 1994.

[15] A. Y. Lee, J. M. Connors, P. Klimo, S. E. O'Reilly, and R. D. Gascoyne. Late relapse in patients with diffuse large-cell lymphoma treated with MACOP-B. *J. Clin. Oncol.*, 15(5):1745–1753, May 1997.

[16] C. Muller, N. Murawski, M. H. Wiesen, G. Held, V. Poeschel, S. Zeynalova, M. Wenger, C. Nickenig, N. Peter, E. Lengfelder, B. Metzner, T. Rixecker, C. Zwick, M. Pfreundschuh, and M. Reiser. The role of sex and weight on rituximab clearance and serum elimination half-life in elderly patients with DLBCL. *Blood*, 119(14):3276–3284, Apr 2012.

[17] N. Murawski and M. Pfreundschuh. New drugs for aggressive B-cell and T-cell lymphomas. *Lancet Oncol.*, 11(11):1074–1085, Nov 2010.

[18] N. Murawski, M. Pfreundschuh, S. Zeynalova, V. Poeschel, M. Hanel, G. Held, N. Schmitz, A. Viardot, C. Schmidt, M. Hallek, M. Witzens-Harig, L. Trumper, T. Rixecker, and C. Zwick. Optimization of rituximab for the treatment of DLBCL

(I): dose-dense rituximab in the DENSE-R-CHOP-14 trial of the DSHNHL. *Ann. Oncol.*, 25(9):1800–1806, Sep 2014.

[19] Houy Nicolas and Le Grand François. Optimizing immune cell therapies with artificial intelligence. *Journal of Theoretical Biology*, 461:34–40, January 2019.

[20] M. Okroj, A. Osterborg, and A. M. Blom. Effector mechanisms of anti-CD20 monoclonal antibodies in B cell malignancies. *Cancer Treat. Rev.*, 39(6):632–639, Oct 2013.

[21] M. Pfreundschuh, V. Poeschel, S. Zeynalova, M. Hanel, G. Held, N. Schmitz, A. Viardot, M. H. Dreyling, M. Hallek, C. Mueller, M. H. Wiesen, M. Witzens-Harig, L. Truemper, U. Keller, T. Rixecker, C. Zwick, and N. Murawski. Optimization of rituximab for the treatment of diffuse large B-cell lymphoma (II): extended rituximab exposure time in the SMARTE-R-CHOP-14 trial of the german high-grade non-Hodgkin lymphoma study group. *J. Clin. Oncol.*, 32(36):4127–4133, Dec 2014.

[22] M. Pfreundschuh, J. Schubert, M. Ziepert, R. Schmits, M. Mohren, E. Lengfelder, M. Reiser, C. Nickenig, M. Clemens, N. Peter, C. Bokemeyer, H. Eimermacher, A. Ho, M. Hoffmann, R. Mertelsmann, L. Trumper, L. Balleisen, R. Liersch, B. Metzner, F. Hartmann, B. Glass, V. Poeschel, N. Schmitz, C. Ruebe, A. C. Feller, and M. Loeffler. Six versus eight cycles of bi-weekly CHOP-14 with or without rituximab in elderly patients with aggressive CD20+ B-cell lymphomas: a randomised controlled trial (RICOVER-60). *Lancet Oncol.*, 9(2):105–116, Feb 2008.

[23] M. Pfreundschuh, L. Trumper, M. Kloess, R. Schmits, A. C. Feller, C. Rube, C. Rudolph, M. Reiser, D. K. Hossfeld, H. Eimermacher, D. Hasenclever, N. Schmitz, and M. Loeffler. Two-weekly or 3-weekly CHOP chemotherapy with or without etoposide for the treatment of elderly patients with aggressive lymphomas: results of the NHL-B2 trial of the DSHNHL. *Blood*, 104(3):634–641, Aug 2004.

[24] M. Pfreundschuh, L. Trumper, M. Kloess, R. Schmits, A. C. Feller, C. Rudolph, M. Reiser, D. K. Hossfeld, B. Metzner, D. Hasenclever, N. Schmitz, B. Glass, C. Rube, and M. Loeffler. Two-weekly or 3-weekly CHOP chemotherapy with or without etoposide for the treatment of young patients with good-prognosis (normal LDH) aggressive lymphomas: results of the NHL-B1 trial of the DSHNHL. *Blood*, 104(3):626–633, Aug 2004.

[25] M. Pfreundschuh, L. Trumper, A. Osterborg, R. Pettengell, M. Trneny, K. Imrie, D. Ma, D. Gill, J. Walewski, P. L. Zinzani, R. Stahel, S. Kvaloy, O. Shpilberg,

U. Jaeger, M. Hansen, T. Lehtinen, A. Lopez-Guillermo, C. Corrado, A. Scheliga, N. Milpied, M. Mendila, M. Rashford, E. Kuhnt, and M. Loeffler. CHOP-like chemotherapy plus rituximab versus CHOP-like chemotherapy alone in young patients with good-prognosis diffuse large-B-cell lymphoma: a randomised controlled trial by the MabThera International Trial (MInT) Group. *Lancet Oncol.*, 7(5):379–391, May 2006.

[26] M. Pfreundschuh, C. Zwick, S. Zeynalova, U. Duhrsen, K. H. Pfluger, T. Vrieling, R. Mesters, H. G. Mergenthaler, H. Einsele, M. Bentz, E. Lengfelder, L. Trumper, C. Rube, N. Schmitz, and M. Loeffler. Dose-escalated CHOEP for the treatment of young patients with aggressive non-Hodgkin's lymphoma: II. Results of the randomized high-CHOEP trial of the German High-Grade Non-Hodgkin's Lymphoma Study Group (DSHNHL). *Ann. Oncol.*, 19(3):545–552, Mar 2008.

[27] G. L. Plosker and D. P. Figgitt. Rituximab: a review of its use in non-Hodgkin's lymphoma and chronic lymphocytic leukaemia. *Drugs*, 63(8):803–843, 2003.

[28] K. Roesch, D. Hasenclever, and M. Scholz. Modelling lymphoma therapy and outcome. *Bull. Math. Biol.*, 76(2):401–430, Feb 2014.

[29] K. Roesch, M. Scholz, and D. Hasenclever. Modeling combined chemo- and immunotherapy of high-grade non-Hodgkin lymphoma. *Leuk. Lymphoma*, 57(7):1697–1708, 07 2016.

[30] M. A. Shipp, D. P. Harrington, M. M. Klatt, M. S. Jochelson, G. S. Pinkus, J. L. Marshall, D. S. Rosenthal, A. T. Skarin, and G. P. Canellos. Identification of major prognostic subgroups of patients with large-cell lymphoma treated with m-BACOD or M-BACOD. *Ann. Intern. Med.*, 104(6):757–765, Jun 1986.

[31] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489, January 2016.

[32] S. H. Swerdlow, E. Campo, S. A. Pileri, N. L. Harris, H. Stein, R. Siebert, R. Advani, M. Ghielmini, G. A. Salles, A. D. Zelenetz, and E. S. Jaffe. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*, 127(20):2375–2390, 05 2016.

[33] T. van Meerten, R. S. van Rijn, S. Hol, A. Hagenbeek, and S. B. Ebeling. Complement-induced cell death by rituximab depends on CD20 expression level and acts complementary to antibody-dependent cellular cytotoxicity. *Clin. Cancer Res.*, 12(13):4027–4035, Jul 2006.

Appendix

A Description of the model

We report below the ODE of the model we use as well as its calibration. We recall that the model borrows from [29] and the calibration from [16, 28, 29]. T is the number of tumor cells and E a surrogate number of immune effector cells.

$$\begin{cases} \dfrac{dT}{dt} &= \alpha T - \nu E T^c - k_T T 1_{CT} - k_{Rit} C_{Rit} T \\ \dfrac{dE}{dt} &= \sigma + \rho(1 + f_{Rit} 1_{Rit}) \dfrac{E T^c}{\eta + T^c} - \delta E - \mu E T^c - k_E E 1_{CT} \end{cases},$$

with

- 1_{CT} equals to 1 if the last chemotherapy application occured less than 1 day before the current date (0 otherwise),
- 1_{Rit} equals to 1 if the last Rituximab application occured less than 100 days before the current date (0 otherwise),
- C_{Rit} is the Rituximab plasmatic concentration. It is the sum, for all administrations, of the output of an open 2-compartment model as given in [16] with Q the intercompartmental clearance, V the central volume, $t_{1/2,\alpha}$ and $t_{1/2,\beta}$ the elimination half-lives.

Parameters' values are given in Table 6.

Parameter	Value	Dimension	Source	Description
σ	1400	cells.day^{-1}	[29]	Tumor-independent production rate of effector cells
η	2.019e5	cells	[29]	number of tumor cells where effector cell stimulation rate is half-maximal
μ	3.422e-10	cell^{-1}day^{-1}	[29]	Tumor-induced inactivation rate of effector cells
δ	0.007	day^{-1}	[29]	inactivation rate of Effector cells
ν	1.101e-7	cell^{-1}day^{-1}	[29]	Effector-induced elimination rate of tumor cells
c	0.75		[29]	Exponent in interaction terms of effector and tumor cells corresponding to dimensionality of tumor surface
k_{Rit}	0.6e-4	day^{-1}ml$^{-1}\mu$g	[29]	Factor of tumor cell kill due to rituximab therapy
f_{Rit}	0.3		[29]	Immune stimulation factor of rituximab
k_E	0.476	day^{-1}	[29]	Effector log cell kill due to chemotherapy
α	$\mathcal{LN}$(0.118,0.054)	day^{-1}	[29]	Tumor growth rate
ρ	$\mathcal{LN}$(0.078,0.038)	day^{-1}	[29]	Tumor-induced stimulation rate of effector cells
$\log_{10}(T_{diag})$	$\mathcal{N}$(11.315,0.473)	cells	[29]	log10 tumor volume at diagnostic
k_T	$\mathcal{N}$(2.363,0.307)	day^{-1}	[29]	Tumor log cell kill due to chemotherapy
Q	$\mathcal{N}$(0.408,0.072)	L/day	[16]	Intercompartmental clearance
V	$\mathcal{N}$(3.88,0.2)	L	[16]	Central volume
$t_{1/2,\alpha}$	0.3529	day	[16]	Elimination half-life
$t_{1/2,\beta}$	37.333	day	[16]	Elimination half-life
BSA	$\mathcal{N}$(1.8,0.15)	m^2	[16]	Body Surface Area

Table 6: Parameters' values.
$\mathcal{LN}$ and $\mathcal{N}$ correspond to a log-normal and a normal distribution, respectively.

B Additional results

B.1 Values of Figure 2

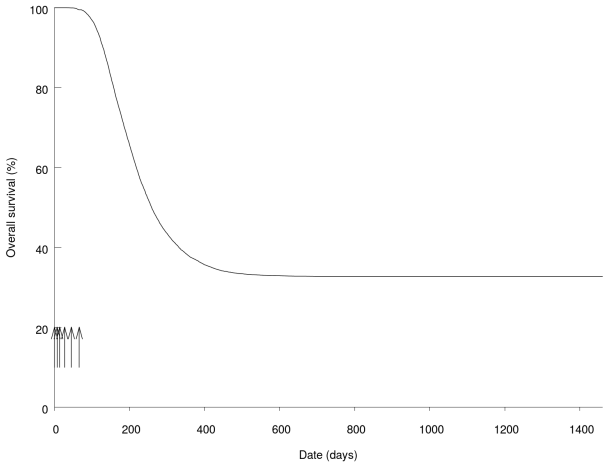
We report in Table 7 the values underlying Figure 2. In particular, we report the period between injections that corresponds to the highest two-year overall survival probability. Note that this period is not defined when there is no or only one injection. We also report the exact values for the corresponding median overall survival probability, as well as the associated confidence interval.

x	y^*	OSP (observed)	OSP (95% CI)
0		0.41	[0.30 - 0.56]
1		4.42	[4.03 - 4.84]
2	7	10.90	[10.30 - 11.53]
3	7	17.61	[16.88 - 18.37]
4	7	23.72	[22.90 - 24.56]
5	8	27.91	[27.04 - 28.80]
6	10	30.75	[29.85 - 31.66]
7	11	32.09	[31.18 - 33.01]
8	12	32.51	[31.60 - 33.43]
9	13	32.44	[31.53 - 33.36]
10	14	32.04	[31.13 - 32.96]
11	15	31.57	[30.67 - 32.49]
12	15	31.33	[30.43 - 32.25]

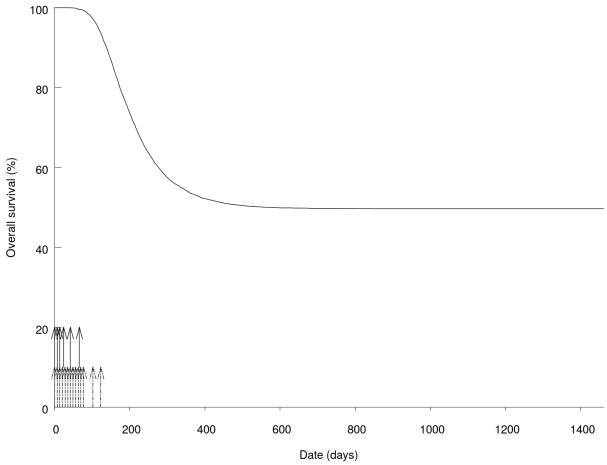
Table 7: CHOP cycles xCHOPy^*, where for each number of cycles x, the cycle length y^* maximizes the two-year overall survival probability (OSP).

B.2 Injection timing of optimal protocols

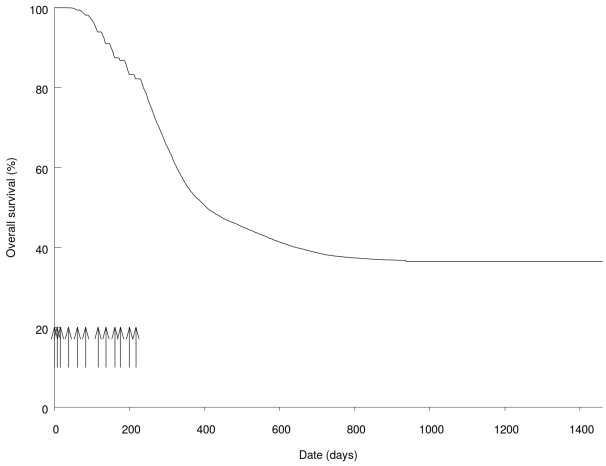
In Figure 5, we plot the evolution of overall survival probabilities and the timing of injections for the optimal protocols $OC_{6,0}$, $OC_{6,14}$, $OC_{12,0}$, $OC_{12,8}$, and $OC_{12,14}$ that are presented in Table 2. They confirm the findings of Figure 3 presented in Section 4.



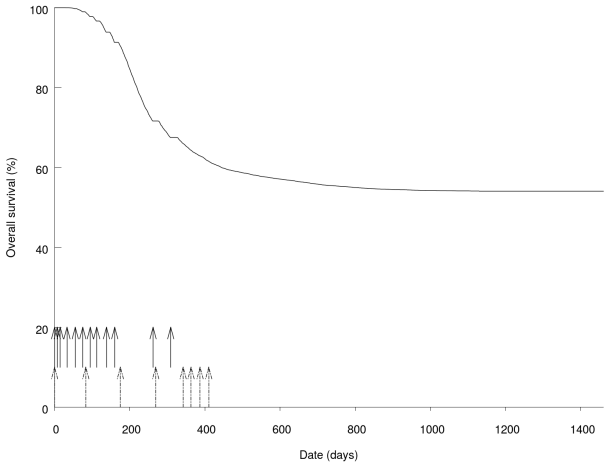
(a) Optimal protocol $OC_{6,0}$.



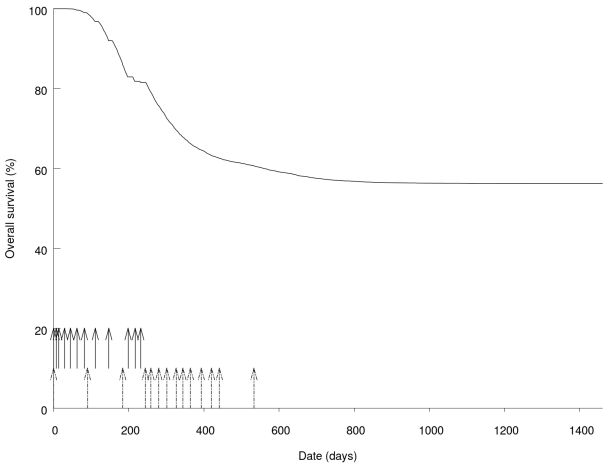
(b) Optimal protocol $OC_{6,14}$.



(c) Optimal protocol $OC_{12,0}$.



(d) Optimal protocol $OC_{12,8}$.



(e) Optimal protocol $OC_{12,14}$.

Figure 5: The black curve represents the evolution of the overall survival probabilities. Arrows correspond to injections: plain ones for chemotherapy and dashed ones for immunotherapy.