Personalized oncology with artificial intelligence: The case of temozolomide

Nicolas Houy* François Le Grand†

July 8, 2019

Running title: Personalized oncology with artificial intelligence.

Abstract

Purpose. Using artificial intelligence techniques, we compute optimal personalized protocols for temozolomide administration in a population of patients with variability.

Methods. Our optimizations are based on a Pharmacokinetics / Pharmacodynamics (PK/PD) model with population variability for temozolomide, inspired by Faivre et al. [10] and Panetta et al. [25, 26]. The patient pharmacokinetic parameters can only be partially observed at admission and are progressively learned by Bayesian inference during treatment. For every patient, we seek to minimize tumor size while avoiding severe toxicity, *i.e.* maintaining an acceptable toxicity level. The optimization algorithm we rely on borrows from the field of artificial intelligence.

Results. Optimal personalized protocols (OPP) achieve a sizable decrease in tumor size at the population level but also patient-wise. The tumor size is on average 67.2 grams lighter than with the standard maximum-tolerated dose protocol (MTD) after 336 days (12 MTD cycles). The corresponding 90% confidence interval for average tumor size reduction amounts to $[58.6 - 82.7]$ (grams). When treated with OPP, less patients experience severe toxicity in comparison to MTD.

Major findings. We quantify *in-silico* the benefits offered by personalized oncology in the case of temozolomide administration. To do so, we compute optimal personalized protocols for a population of heterogeneous patients using artificial intelligence techniques. At each treatment day, the protocol is updated by taking into account the feedback obtained from patient's reaction to the drug administration. Personalized protocols greatly differ from each other, and from the standard MTD protocol. Benefits of personalization are very sizable: tumor sizes are much smaller on average and also patient-wise, while severe toxicity is made less frequent.

*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.

Keywords: Pharmacokinetics, Pharmacodynamics, Optimization, Personalized Oncology, Artificial Intelligence.

Conflicts of interest: Both authors have no conflict of interest to declare.

Quick guide to equations and assumptions

We consider a PK/PD model built on Faivre et al. [10] and Panetta et al. [25, 26].

We assume that patients are endowed with heterogeneous pharmacokinetic (PK) parameters. More precisely, since the PK part of the model is driven by a one-compartment model, values for the volume of distribution V, the constant of absorption k_a and the constant of elimination k_e are different from one patient to another. In addition, these values cannot be perfectly observed and therefore cannot directly be used in the process of determining optimal personalized protocols.

For each patient, we conduct a personalized optimization to determine the protocol maximizing the efficacy –i.e., yielding the smallest tumor size after 336 days (12 MTD cycle length)– while maintaining an acceptable toxicity level. We proceed as follows.

- To handle the dimensionality of individual optimization problems, we use an algorithm borrowing from the class of Monte-Carlo tree search in artificial intelligence. See [15] for an application to the determination of optimal protocol for a population of heterogeneous patients (same protocol for all).

- PK parameters are observed with noise. The algorithm can only rely on beliefs over PK parameter values. Formally, these beliefs are probability distributions over possible values.

- At the beginning of the protocol, the algorithm relies on initial beliefs, that only depend on the patient's body surface area (BSA) and are therefore not highly personalized.

- At each treatment day, new PK information is available, that is used to update beliefs about patient's PK parameters by Bayesian inference.

- Because initial beliefs only depend on BSA, protocol personalization is not very pronounced in the first days. However, due the successive Bayesian updates, heterogeneity among protocols rises with time and eventually becomes sizable.

1 Introduction

The term "personalized medicine" has been coined to describe the set of techniques enabling to administer the right drug to the right disease and the right patient. In oncology, this expression covers a wide range of fields and methods promising to tailor the treatment to the patient and her tumor [4]. Pharmacogenetics and pharmacokinetics are two promising fields for personalized oncology [6]. Pharmacogenetics involves genetic profiling both at the patient and at the tumor levels. At the patient level, pharmacogenomic predictors have been developed for selecting drugs before administration [21]. For instance, a polymorphism affecting gene UGT1A1 has been proved to drive the toxicity of irinotecan, which is a common drug for colorectal cancer [16]. US FDA recommends to test for the presence of the incriminated allele before irinotecan administration. However, properly assessing the role of a genetic biomarker is a difficult task and does not always offer a clear-cut result. While polymorphisms of the cytochrome CYP2D6 have been thought, for approximately 10 years, to affect the efficacy of tamoxifen –used in the treatment of breast cancers–, this relationship is now much less consensual [17]. For this reason, the most promising route for pharmacogenetics appears to be the genetic profiling of cancer tumors [13]. For instance, the treatment of non-small-cell lung cancers has been shown to benefit from the administration of first-line tyrosine kinase inhibitors provided that tumors exhibit particular mutations on the EGFR or ALK genes [22]. Consequently, the FDA has now added recommendations for a genetic testing in a number of package inserts of anticancer drugs. Though highly promising, genetic tumor testing raises a number of issues. In particular, this is an invasive technique requiring costly and complex methods. Furthermore, the very unstable nature of the tumor and of its genetic profile means that the conclusions drawn at a given point in time may not hold later on.

Besides pharmacogenetics, recent progress has also been accomplished in pharmacokinetics and pharmacodynamics, notably supported by mathematical and computational tools. Indeed, a number of methods have been used to personalize cancer protocols and thereby aim at a better efficacy-toxicity trade-off [1]. First, regression techniques have been developed in population pharmacokinetics to determine how patient characteristics affect the relationship between drug doses on the one side and efficacy and toxicity on the other side (see [28] for a review). This approach has enabled to identify patient covariates, such as for instance body surface area (BSA, henceforth), weight, sex or age, that are useful tools for oncologists to personalize protocols in everyday practice. Second, probabilistic models of toxicity have been used to guide drug dosing [7]. Third, optimizations of PK/PD models have also helped determine –mostly in *in-silico* trials– both the optimal dosing and the optimal schedule. This is for instance the case of vinorelbine. While the standard metronomic protocol [8] relies on the administration of 50 mg on days 1, 3, and 5, Barbolosi and colleagues [5] show that administering 60 mg, 30 mg and 60 mg on days 1, 2, and 4 yields a smaller tumor mass, without aggravating toxicity. Such PK/PD optimizations have also

been used to guide the determination of protocols in phase I/II clinical trials. MODEL1 study [14, 23] is a pioneering example of such trials. Relatedly, *in-silico* trials relying on mathematical models and computational simulations have also been used for personalizing cancer immunotherapy (see [18] for a seminal reference on mathematical modelling in immunotherapy). Indeed, in spite of a number of successes, immunotherapy outcomes significantly vary from one patient to another. Properly calibrated mathematical models can prove to be helpful for personalizing therapy and improving its results (see [2] for a review and [19, 20] for the application of mathematical models to guide immunotherapy personalization). Interestingly, despite involving different therapies, mathematical models for chemotherapy and immunotherapy share a number of common characteristics (potential benefits of *in-silico* trials, the aim of treatment personalization for instance) and, relatedly, similar difficulties, such as those related to personalized clinical trials (see [3] in the context of immunotherapy).

This paper aims at quantifying *in-silico* the benefits of optimal personalized protocols in the case of temozolomide. Temozolomide is prescribed in the treatment of some brain cancers, notably for children. The main difference with previous works is that we determine the best *individual* protocol for each patient. Our model assumes patient heterogeneity in the population, and that patient characteristics cannot be perfectly observed. In consequence, for each patient, the optimization relies on beliefs about patient characteristics that are revised by Bayesian update at every information arrival. At each date, we determine for each patient the best possible treatment decision – including no treatment – given the knowledge about the patient characteristics at that date. Our optimization objective consists in maximizing efficacy, while maintaining an acceptable toxicity level. Our measure for efficacy is the tumor size at the end of our 336-day protocol period and we consider toxicity as acceptable, when the minimal normalized absolute neutrophil count (ANC nadir, henceforth) over the protocol period never goes below a given severe toxicity threshold.

The main lesson from the optimal personalized protocols (OPP, henceforth) we compute is that they offer very sizable gains in terms of efficacy with less frequent severe toxicity. Our benchmark protocol is the standard Maximum Tolerated Dose (MTD) protocol, which consists in a 28-day cycle with the administration of 200 mg/m^2 from day 1 to day 5. *In-silico* experiments enable us to administer two different protocols to two populations having the exact same characteristics. Control and treatment groups are perfectly identical and outcome differences purely reflect differences in protocol effects and not differences in population characteristics. Our experiment concerns two strictly identical populations of 192 heterogeneous patients to which we administer either MTD or OPP. In terms of efficacy, the tumor size is on average 67.2 grams lighter with OPP than with MTD and the 90% confidence interval for average tumor size reduction is $[58.6 - 82.7]$ (grams). Tumor size gains are not homogeneous among the patient population and mostly stem from patients

with a very high tumor mass when MTD is administered. Consequently, tumor mass distribution is much less dispersed with OPP than with MTD. In our population, MTD leads to a 90%-confidence interval of $[0.001 - 292.9]$ (grams) for tumor mass, while the same interval reduces to $[10^{-5} - 61.5]$ with OPP. This strong reduction in tumor masses is not accompanied by more severe toxicity. Indeed, only 3 patients in the population experience an ANC nadir below the acceptable threshold with OPP, while this number amounts to 10 with MTD.

We do not impose any cycle constraint on personalized protocols, assuming that drug administration can occur at any day of the week, week-end or day-off included. We believe that imposing such cycle constraints would not have been consistent with a full-fledged protocol personalization, that should offer the highest possible flexibility for maximizing patient benefits. A consequence of this absence of cycle constraint is that personalization generates very heterogeneous protocols that greatly differ from one patient to another. More precisely, most protocols share a common treatment pattern for the first few days, which are used to learn about patient reaction to drug and about patient characteristics. After these first few days, the heterogeneity among protocols starts to grow and becomes increasingly large with time. At one side of the distribution, some patients only receive a couple of doses after the initial treatment period. On the other side, other patients face a very intense administration schedule, which involves a number of doses much higher than MTD.

Finally, our work also offers a methodological innovation since we rely on artificial intelligence algorithms to solve the optimization problems. We already used a similar technique to compute the optimal protocol – without personalization or noise on the observation of patient characteristics – for a population of patients [15]. Here, we extend the optimization algorithms to the determination of personalized protocol for patients whose characteristics are imperfectly observed and progressively learned by Bayesian inference. Our results show that the combination of PK/PD models with artificial intelligence techniques offers promising perspectives for the development of optimal personalized protocols in oncology.

2 Materials and methods

2.1 PK/PD model

We use a PK/PD model for temozolomide that consists of three parts. The first two parts dealing with pharmacokinetics and pharmacodynamics for efficacy rely on Faivre et al. [10]. The third part about pharmacodynamics for toxicity borrows from Panetta and coworkers [26].[1]

Let us briefly describe the three blocks of the model. First, a standard one-compartment

[1] A detailed presentation of the model we use can be found in Section 1 of the supplemental material.

model, initially proposed by Panetta et al. [25], drives the pharmacokinetics. This model features population variability in pharmacokinetic parameters, which are therefore patient-dependent. We also assume that these pharmacokinetic parameters can only be imperfectly observed. Second, we use the model of [10] – which is itself based on the interface model of Meille et al. [24] – for pharmacodynamics of efficacy. Temozolomide is assumed to affect in a different way cancer cells and endothelial ones. The latter are supposed to be less sensitive than the former to the drug, but they do not feature drug resistance over time. This model simultaneously captures standard cytotoxic effects featuring drug resistance and anti-angiogenic effects. Finally, we follow Panetta and coworkers [26] for modelling pharmacodynamics of toxicity using a physiological model of hematopoiesis. The model mimics the different development stages of proliferating cells in the bone marrow, from pluripotential stem cells to differentiated blood cells. Temozolomide affects toxicity through a very direct channel, as temozolomide simply shuts down the growth of proliferating cells in the bone marrow, which eventually hurts neutrophil counts.

2.2 Simulations

Our simulation period covers a time length of 336 days, which corresponds to 12 full cycles of 28 days for the standard MTD protocol. All computations are implemented in C++.

We consider a population of 192 heterogeneous patients, who differ in their pharmacokinetic parameters –that cannot be directly observed by the practitioner. At day 0, all patients display a 30 grams tumor. We administer different protocols to populations with the exact same characteristics, such that differences in outcomes are only due to differences in protocols but not to differences in populations. One of the advantages of *in-silico* experiments is to allow for exactly identical control and treatment groups, which enables us to perfectly isolate the protocol effect. Of note, the population size is sufficiently large to generate statistically significant differences among the protocols we consider.

For each protocol under consideration and each patient in the population, we assess the efficacy and toxicity as follows. We measure efficacy by the tumor size (in grams) at the final day of the 336-day period. We measure toxicity by the minimal normalized ANC (in %) –or normalized ANC nadir– obtained over the simulation period of 336 days. Toxicity will be considered to be acceptable when the ANC nadir does not cross a given threshold. We will define this threshold as the ANC nadir reached by the 5^{th} percentile of toxicity in a population treated with MTD: 2.59%. We have made the choice to measure toxicity by the ANC nadir only, since it is an unambiguous measure of toxicity and for instance appears in prescribing information [11]. However, this choice does not result from a limitation of the algorithm, which could handle other measures of toxicity just as well (as in [14, 23]). It is noteworthy that our choice of the toxicity level of 2.59% is arbitrary, but is not key to explain our results. Indeed, this value – that corresponds by definition to

the 5^{th} percentile of toxicity level in a population treated with MTD – is used as an input of the toxicity constraint in OPP. If we had chosen another value for the limit threshold of severe toxicity, then the results of OPP would have been modified accordingly and the magnitude of the advantage of OPP over MTD would have remained of similar magnitude. The same conclusion would have held if we had used a different severe toxicity threshold definition. Such an alternative definitive could for instance account for MTD treatment halt because of low tolerance to the treatment by some individuals, [12].

2.3 Optimization and personalization

We conduct personalized protocol optimization in two steps. First, before the beginning of the treatment, we adapt the doses to the BSA of the patient. We follow Panetta and coauthors [25], who show that BSA is the only major covariate affecting the absorption of temozolomide in infants and children with primary central nervous system tumors. This is the first personalization step, to which we will refer as *static* since it is never updated during the protocol period, no matter the patient's reaction to the treatment.

The second step is the dynamic personalized optimization.[2] For each patient, we seek to determine the protocol which maximizes the efficacy, while maintaining an acceptable toxicity level. At each day of our protocol duration of 336 days, we determine which dose is optimal, given the current patient state and the information available at that day. The PK/PD model for temozolomide enables us to simplify the dosing possibilities and to restrict our attention to a choice between the maximum tolerated dose of 200 mg/m^2 and no dose. Indeed, temozolomide doses lower than 200 mg/m^2 imply a similar toxicity level but a lower efficacy. This stems from three features of the temozolomide PK/PD model: (i) the growth shut-down for proliferating cells appears for relatively low plasmatic concentrations, and (ii) the plasmatic clearing of temozolomide is quite fast, while (iii) temozolomide becomes really effective only at large concentrations.

The dynamic optimization we conduct is complex for two main reasons. First, the size of the problem implies that it cannot be handled with standard optimization techniques, such as dynamic programming. Indeed, there are more than 10^{100} (a googol!) possible treatments to be tested for each patient of the population. Even *in-silico* experiments are physically unable to investigate such a tremendous number of possibilities. To circumvent this difficulty, we use an algorithm belonging to the class of Monte-Carlo tree search in artificial intelligence and for instance used in the famous Go program AlphaGo [9, 27]. We have already applied a similar algorithm to determine the optimal protocol (without personalization) for a population with patient variability in [15].

The second difficulty in our optimization is the imperfect observation of patient pharmacokinetic parameters. For each patient, the algorithm can only rely on beliefs about

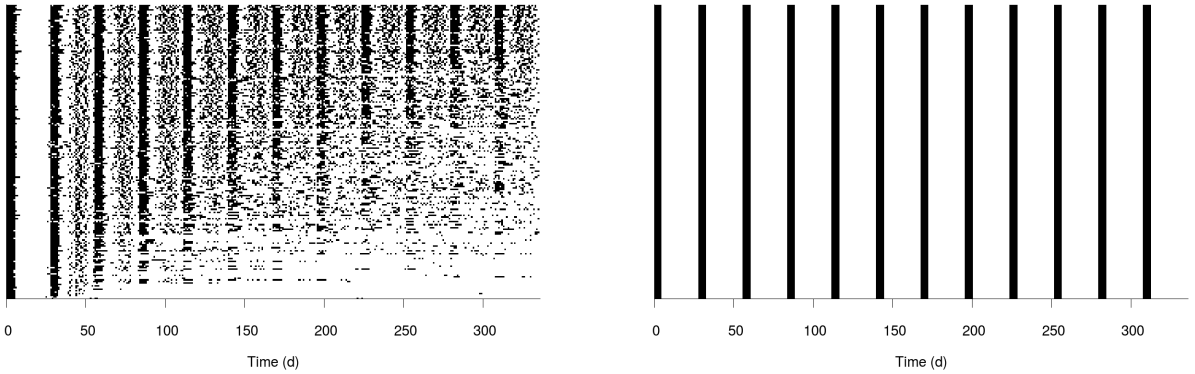
[2]We present in detail the algorithm determining optimal personalized protocols in Section 2 of the supplemental material.

patient pharmacokinetics, and not on actual values. Formally, these beliefs are probability distributions over patient pharmacokinetics. At the beginning of the protocol, there is no specific information about the patient, except her BSA. In consequence, the protocol is not highly personalized in the first days of the protocol. In addition to the observation of the BSA, a second source of information is progressively made available for each patient: the way she reacts to the treatment. Hence, after a few days of treatment, as new information about the patient has arrived, beliefs about patient pharmacokinetics are updated by Bayesian inference and taken into account in the optimization. In consequence, protocols become increasingly personalized over time. Because of these regular updates in the protocol optimization, we will refer to it as *dynamic* personalization.

Note that we suppose that new information about patients' characteristics comes at every treatment day. Indeed, this information acquisition is not very invasive and can be inferred from blood samples with the patient already in the hospital facility.

3 Results

We compare our optimal personalized protocols (OPP) to the standard MTD. Of note, MTD embeds neither optimization, nor personalization (be it static or dynamic).



(a) OPP (b) MTD

Figure 1: Time evolution of protocols for the population. Every row represents a patient and every column a day. A black square is a treatment day and a white one a rest day.

We start with illustrating the protocol schedules for the two populations to which either MTD or OPP has been administered. In Figure 1, we plot treatment and rest days for MTD and OPP. On each panel, every row corresponds to a different patient and every column to a day of the protocol period. A black square represents a treatment day and a white one a rest day. Unsurprisingly, panel 1b for MTD displays regularly spaced vertical lines. This is the graphical illustration that (i) the same protocol is administered to all

patients and that (ii) the protocol features a fixed 28-day cycle. Panel 1a confirms that OPP become increasingly personalized with time and that the overall heterogeneity among protocols is very sizable. Some protocols only include a handful of treatment days, while others feature a number of treatment days higher than MTD.

We now turn to the outcomes of the protocols in terms of efficacy and toxicity. Population-wide results are reported in Table 1.

Protocol	Tumor mass (g)	Number of patients with severe toxicity	Norm. ANC nadir (%)
MTD	79.86 [0.001−292.9]	10 [3.14%−8.52%]	6.62 [2.59−10.76]
OPP	12.68 [0.00001−61.46]	3 [0.63%−3.84%]	2.97 [2.65−3.66]

Table 1: MTD and OPP protocols comparison. 192 patients.
Tumor mass and normalized ANC nadir: Median values and in square brackets, the 5^{th} and 95^{th} percentiles.
Number of patients with severe toxicity: observation and in square brackets, 90% confidence interval computed as Wilson score interval.

We observe that OPP enable to significantly reduce tumor sizes with an improvement for toxicity. Tumor mass diminishes by 67.2 grams on average when switching from MTD to OPP. Tumor mass of the 95^{th} percentile decreases by more than 230 grams. OPP offer sizable efficacy gains on average, and these gains are especially large for patients for whom MTD has a poor efficacy and yields a high tumor mass. Furthermore, the better efficacy is not accompanied by an aggravation of toxicity, since a fewer number of patients experience severe toxicity when OPP is administered as compared to MTD. This number amounts to 3 patients with OPP and 10 patients with MTD, while both population sizes amount to 192. Similarly, the 5^{th} percentile of ANC nadir implied by OPP is 2.65%, which is an improvement over the 2.59% implied by MTD. Of note, the normalized ANC nadir with OPP is less dispersed than with MTD. With OPP, the 95^{th} percentile for ANC nadir amounts to 3.66% while it is 10.76% with MTD. The better efficacy of OPP as compared to MTD can partly be explained by a better management of toxicity at the patient level. Indeed, OPP manages to keep the tumor size low by decreasing the normalized ANC nadir for most individuals, yet keeping it in the toxicity range allowed by the imposed constraint and at no cost given the objective. We recall that the the objective of OPP is to minimize the tumor size, subject to a unique constraint, which is to avoid crossing the 2.59% normalized ANC threshold. Of note, we could include a number of other constraints, and consider other optimization objectives. One of the main strength of the OPP algorithm we use – and that is based on Monte Carlo tree search algorithms – is its versatility. Accounting for supplementary constraints or different objectives would therefore be painless. We leave the investigation of these alternative constraints and objectives for further research.

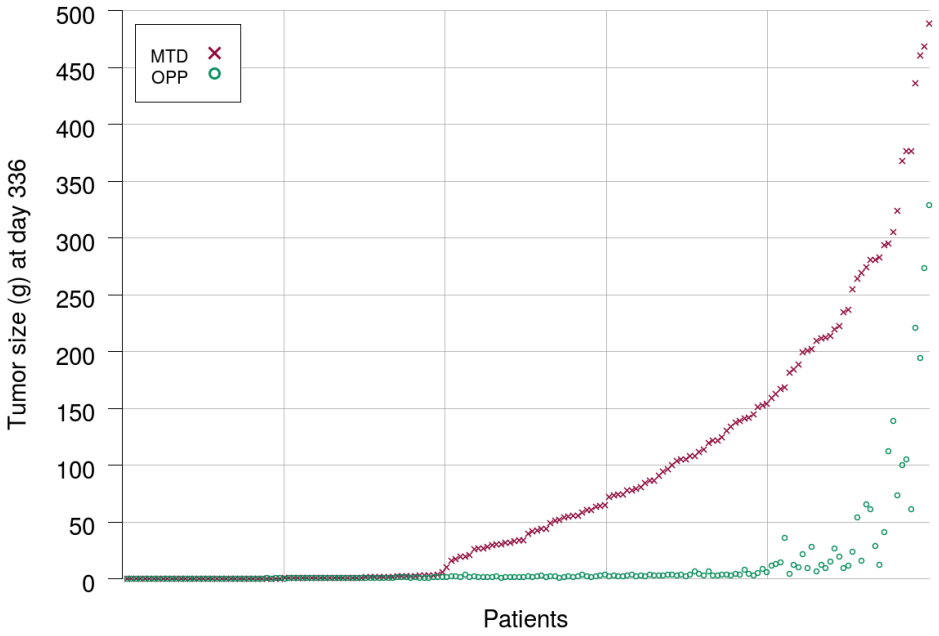


Figure 2: Patient-wise comparison of protocols. Each vertical line displays the tumor size for two patients with identical characteristics, who are administered either MTD (red cross) or OPP (green circle).

Figure 2 plots tumor sizes at the patient level for both MTD and OPP. Each vertical line represents patients with the exact same characteristics to whom either MTD or OPP has been administered. The graph shows that gains are especially large when tumor mass remains large with MTD administration. In fact, even though results are not very visible for low tumor sizes, most patients experience a lower tumor mass with OPP than with MTD. This aspect is more clearly illustrated by Table 2.

	Proportions of patients (%)	Average tumor size difference (grams)
OPP better than MTD	92.18	76.64
MTD better than OPP	7.82	0.09

Table 2: MTD and OPP protocols comparison conditionnally on best protocol.

OPP has a better efficacy than MTD for 92.18% of patients in the population and the tumor size is then on average 76.64 grams lighter with OPP than MTD. For the remaining 7.82% of patients, MTD has a better efficacy than OPP, but the tumor size improvement in those cases is very low and amounts to 0.09 gram only. In conclusion, there are very few patients for which MTD does better than OPP, and when it does, the tumor size difference

is very low. Similarly, we report in Table 3 the descriptive statistics of the difference in tumor sizes (in grams) for the 192 individuals treated with MTD and OPP (the difference is OPP minus MTD).

Individual	Tumor size difference OPP−MTD (grams)
5th percentile	− 0.15
25th percentile	0.03
Median	29.92
75th percentile	118.68
95th percentile	225.39
Average	67.24

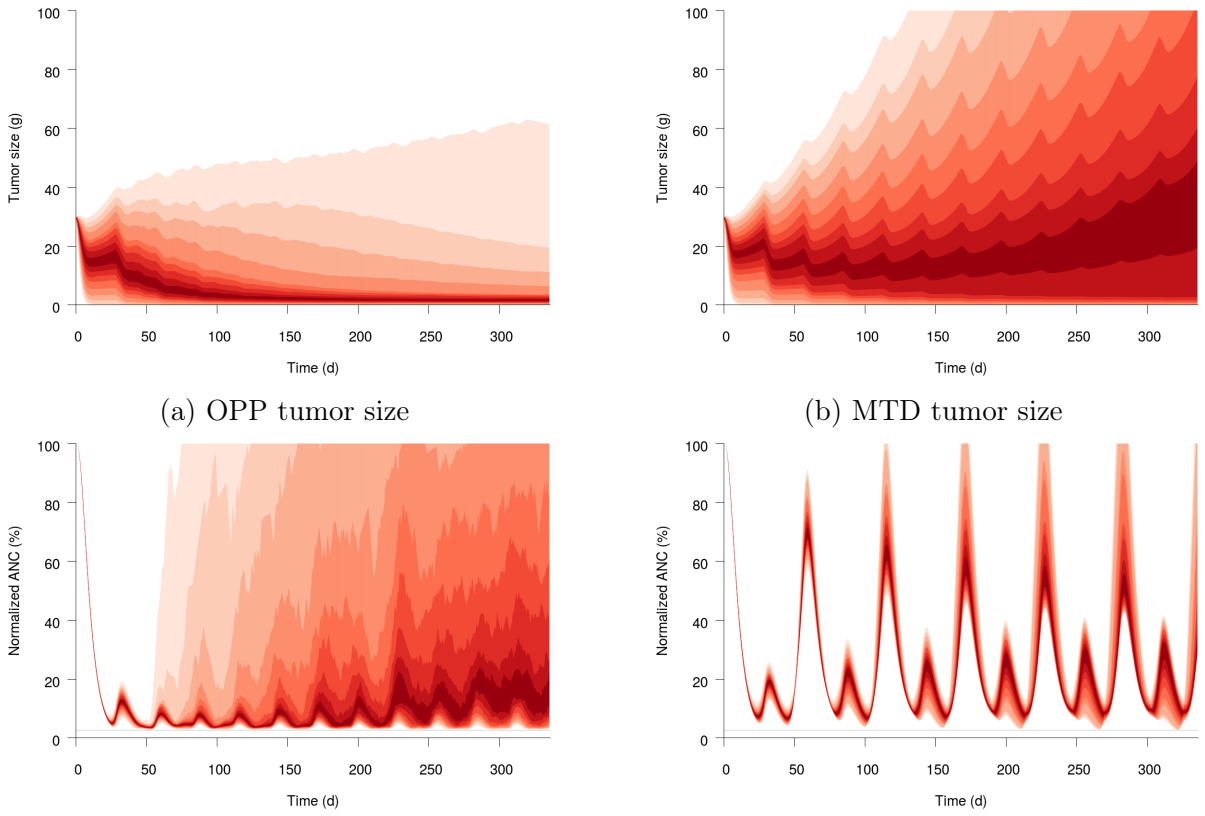
Table 3: Descriptive statistics for the difference in tumor sizes (OPP minus MTD, in grams) for the 192 individuals treated with MTD and OPP.

Finally, Figure 3 gathers the plots of the distribution of temporal evolution for tumor sizes with MTD and OPP (panels 3a and 3b) and for normalized ANC with MTD and OPP (panels 3c and 3d). For each graph, each shade corresponds to a different interval of the distribution. The darkest shade includes the median.[3] The comparison of panels 3a and 3b confirms than OPP significantly reduce tumor sizes on average and also strongly reduce the dispersion of after-treatment tumor sizes. On panel 3c, we observe that the distribution of ANC is highly concentrated at the beginning of the treatment with OPP, but much less afterwards. The higher dispersion with time of ANC in OPP reflects the personalization of protocols and the heterogeneity in treatment doses and schedules.

4 Discussion

Our results show that optimal personalized protocols achieve a remarkable efficacy compared to MTD, as well as an improvement for toxicity. Of note, MTD and OPP are very different protocols: MTD is a non-optimal one-size-fits-all protocol, while OPP correspond to optimal patient-specific protocols. Their different outcomes may therefore reflect several possible factors: (i) the population-wide optimization (with no personalization), (ii) the static personalization by BSA and (iii) the dynamic personalization by Bayesian update of patient characteristics. To isolate the contributions of these three different factors, we compute the outcomes for two other protocols: (i) a population-wide optimal protocol with no personalization, and (ii) optimal individual protocols with BSA personalization (but no dynamic optimization). Results are reported in Table 4 (values for MTD and OPP are repeated).

[3]Formally, each shade accounts for $1/17 \approx 5.88\%$ of the total distribution and the central shade to the



(a) OPP tumor size

(b) MTD tumor size

(c) OPP ANC

(d) MTD ANC

Figure 3: Distribution of the time evolution for tumor sizes (top) and normalized ANC (bottom) for MTD and OPP protocols. Each shade represents a different interval of the distribution.

Table 4 shows that the two most important steps for efficacy improvement over MTD are optimization and dynamic personalization. Compared to these, BSA-personalization has a relatively minor effect. It is interesting to contrast the relative merits of optimization and dynamic personalization. Optimization reduces the tumor mass rather uniformly in the population, as shown by the sharp decline, compared to MTD, in the median tumor mass and the 95^{th} percentile. Dynamic personalization has a more selective impact. On the one hand, OPP has a strong impact on average, that comes from the tumor mass reduction for patients who still had a high tumor mass when the optimal protocol was administered. This is illustrated by the pronounced decline in the 95^{th} percentile of tumor mass. On the other hand, for other patients (loosely speaking, the bottom 50% of the tumor size distribution), one-size-fits-all optimal protocols already offer a high efficacy and yield low tumor masses that are difficult to further reduce. This is reflected in the median tumor

interval $[47.06\% - 52.94\%]$, which includes the median.

Protocol	Average tumor mass (g)	Median tumor mass (g)	Number of patients with severe toxicity
MTD	79.86	31.51 [0.001−292.9]	10 [3.14%−8.52%]
Population-wide optim.	35.53	1.87 [0.001−166.12]	3 [0.63%−3.84%]
BSA-personalized optim.	32.35	1.84 [0.0001−155.7]	6 [1.62%−5.93%]
OPP	12.64	1.74 [0.00001−61.46]	3 [0.63%−3.84%]

Table 4: Protocols comparison. 192 patients.
Median tumor mass, in square brackets, the 5^{th} and 95^{th} percentiles.
Number of patients with severe toxicity: observation and in square brackets, 90% confidence interval computed as Wilson score interval.

size that barely diminishes with OPP compared to non-personalized optimization.

Our paper proposes a first *in-silico* quantification of protocol personalization in the context of temozolomide. Our conclusions are unequivocal: personalization matters and brings significant gains on top of population-wide optimization, especially for patients for whom the one-size-fits-all protocol is not very effective. Obviously, in the current article, many assumptions we made could be discussed with strong arguments. Even though an extensive robustness check is beyond the scope of the present study, we will briefly discuss two aspects. First, our optimization exercise involves a comprehensive exploration of the PK/PD model, while the model has only been calibrated "locally" using a handful of clinical trials that in general only involve a small number of doses and of schedules. Any real-life application would therefore require further tests to check the validity of the model. Of note, this limit is not specific to our paper, and is valid for any optimization. Second, we make a normality assumption on the distribution of patients' PK parameters. This assumption slightly simplifies the learning process and lightens the computational burden. However, for real-life applications, it would be needed either to relax this assumption or at least to check its implications on optimal protocols.

In this article, we worked with an algorithm adapted from the Monte Carlo tree search family and comprehensively described in [15]. We are not claiming that this algorithm is the only one we could implement or even that it is the most efficient if compared to all other known heuristics. We chose to use such an algorithm for the following reasons. First, it takes advantage of the sequential and finite time features of the problem we are facing. Second, it does not require an evaluation function that would be tricky to design in the present framework with high option values of treatment doses. Finally, our algorithm offers a very high degree of flexibility. This feature is particularly interesting for future

implementation in an applied context. A first illustration could consist in integrating decisions regarding complementary exams that are invasive but informative – for instance, if or when a biopsy should be performed depending on the expected informational benefits. A second illustration is related to clinical feasibility. Several elements are indeed likely to affect protocol administration and could be integrated in our optimization algorithm. We can for instance mention deterministic constraints (hospital facilities closed during week-ends or some days-off), or stochastic ones (possible absence of patients on a treatment day). We have chosen not to integrate such constraints in our study, but doing so is rather straightforward. This will obviously lead to slightly less efficient protocols. The magnitude of the loss remains to be quantified.

As a conclusion, we can mention that personalized protocols, as we propose here, are likely to raise a number of practical issues – which are not specific to our method and are common to any personalization device. First, hospital facilities will need to manage a high number of aperiodic patients' schedules, which makes resource allocations (in particular medical and paramedical staff) more difficult and could possibly create shortages if a large number patients happen to need administration on the same day. Second, patients may have a hard time understanding the justifications for their protocol, for instance when talking to other patients, who have the same disease but a very different protocol. This could matter for patients' adherence to protocol administration.

References

[1] Z. Agur, M. Elishmereni, and Y. Kheifetz. Personalizing oncology treatments by predicting drug efficacy, side-effects, and improved therapy: mathematics, statistics, and their integration. *Wiley Interdiscip Rev Syst Biol Med*, 6(3):239–253, 2014.

[2] Z. Agur, K. Halevi-Tobias, Y. Kogan, and O. Shlagman. Employing dynamical computational models for personalizing cancer immunotherapy. *Expert Opin Biol Ther*, 16(11):1373–1385, Nov 2016.

[3] Z. Agur and S. Vuk-Pavlović. Mathematical modeling in immunotherapy of cancer: personalizing clinical trials. *Mol. Ther.*, 20(1):1–2, Jan 2012.

[4] F. Andre, J. Ciccolini, J. P. Spano, F. Penault-Llorca, N. Mounier, G. Freyer, J. Y. Blay, and G. Milano. Personalized medicine in oncology: where have we come from and where are we going? *Pharmacogenomics*, 14(8):931–939, Jun 2013.

[5] D. Barbolosi, J. Ciccolini, C. Meille, X. Elharrar, C. Faivre, B. Lacarelle, N. Andre, and F. Barlesi. Metronomics chemotherapy: time for computational decision support. *Cancer Chemother. Pharmacol.*, 74(3):647–652, Sep 2014.

[6] S. Benzekry, E. Pasquier, D. Barbolosi, B. Lacarelle, F. Barlesi, N. Andre, and J. Ciccolini. Metronomic reloaded: Theoretical models bringing chemotherapy into the era of precision medicine. *Semin. Cancer Biol.*, 35:53–61, Dec 2015.

[7] T. M. Braun, P. F. Thall, H. Nguyen, and M. de Lima. Simultaneously optimizing dose and schedule of a new cytotoxic agent. *Clin Trials*, 4(2):113–124, 2007.

[8] E. Briasoulis, P. Pappas, C. Puozzo, C. Tolis, G. Fountzilas, U. Dafni, M. Marselos, and N. Pavlidis. Dose-ranging study of metronomic oral vinorelbine in patients with advanced refractory cancer. *Clin. Cancer Res.*, 15(20):6454–6461, Oct 2009.

[9] 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.

[10] C. Faivre, D. Barbolosi, E. Pasquier, and N. Andre. A mathematical model for the administration of temozolomide: comparative analysis of conventional and metronomic chemotherapy regimens. *Cancer Chemother. Pharmacol.*, 71(4):1013–1019, Apr 2013.

[11] FDA Approval for Temozolomide. `https://www.accessdata.fda.gov/drugsatfda_docs/label/2006/021029s012lbl.pdf`, 2006.

[12] M. R. Gilbert, M. Wang, K. D. Aldape, R. Stupp, M. E. Hegi, K. A. Jaeckle, T. S. Armstrong, J. S. Wefel, M. Won, D. T. Blumenthal, A. Mahajan, C. J. Schultz, S. Erridge, B. Baumert, K. I. Hopkins, T. Tzuk-Shina, P. D. Brown, A. Chakravarti, W. J. Curran, and M. P. Mehta. Dose-dense temozolomide for newly diagnosed glioblastoma: a randomized phase III clinical trial. *J. Clin. Oncol.*, 31(32):4085–4091, Nov 2013.

[13] D. Hanahan and R. A. Weinberg. Hallmarks of cancer: the next generation. *Cell*, 144(5):646–674, Mar 2011.

[14] E. Henin, C. Meille, D. Barbolosi, B. You, J. Guitton, A. Iliadis, and G. Freyer. Revisiting dosing regimen using PK/PD modeling: the MODEL1 phase I/II trial of docetaxel plus epirubicin in metastatic breast cancer patients. *Breast Cancer Res. Treat.*, 156(2):331–341, Apr 2016.

[15] N. Houy and F. Le Grand. Optimal dynamic regimens with artificial intelligence: The case of temozolomide. *PLoS ONE*, 13(6):e0199076, 2018.

[16] L. Iyer, S. Das, L. Janisch, M. Wen, J. Ramirez, T. Karrison, G. F. Fleming, E. E. Vokes, R. L. Schilsky, and M. J. Ratain. UGT1A1*28 polymorphism as a determinant of irinotecan disposition and toxicity. *Pharmacogenomics J.*, 2(1):43–47, 2002.

[17] C. M. Kelly and K. I. Pritchard. CYP2D6 genotype as a marker for benefit of adjuvant tamoxifen in postmenopausal women: lessons learned. *J. Natl. Cancer Inst.*, 104(6):427–428, Mar 2012.

[18] D. Kirschner and J. C. Panetta. Modeling immunotherapy of the tumor-immune interaction. *J Math Biol*, 37(3):235–252, Sep 1998.

[19] Y. Kogan, K. Halevi-Tobias, M. Elishmereni, S. Vuk-Pavlović, and Z. Agur. Reconsidering the paradigm of cancer immunotherapy by computationally aided real-time personalization. *Cancer Res.*, 72(9):2218–2227, May 2012.

[20] Yuri Kogan, Urszula Foryś, Ofir Shukron, Natalie Kronik, and Zvia Agur. Cellular immunotherapy for high grade gliomas: Mathematical analysis deriving efficacious infusion rates based on patient requirements. *SIAM Journal on Applied Mathematics*, 70(6):1953–1976, 2010.

[21] S. Y. Lee and H. L. McLeod. Pharmacogenetic tests in cancer chemotherapy: what physicians should know for clinical application. *J. Pathol.*, 223(1):15–27, Jan 2011.

[22] T. Li, H. J. Kung, P. C. Mack, and D. R. Gandara. Genotyping and genomic profiling of non-small-cell lung cancer: implications for current and future therapies. *J. Clin. Oncol.*, 31(8):1039–1049, Mar 2013.

[23] C. Meille, D. Barbolosi, J. Ciccolini, G. Freyer, and A. Iliadis. Revisiting Dosing Regimen Using Pharmacokinetic/Pharmacodynamic Mathematical Modeling: Densification and Intensification of Combination Cancer Therapy. *Clin Pharmacokinet*, 55(8):1015–1025, Aug 2016.

[24] C. Meille, A. Iliadis, D. Barbolosi, N. Frances, and G. Freyer. An interface model for dosage adjustment connects hematotoxicity to pharmacokinetics. *J. Pharmacokinet. Pharmacodyn.*, 35(6):619–633, Dec 2008.

[25] J. C. Panetta, M. N. Kirstein, A. Gajjar, G. Nair, M. Fouladi, R. L. Heideman, M. Wilkinson, and C. F. Stewart. Population pharmacokinetics of temozolomide and metabolites in infants and children with primary central nervous system tumors. *Cancer Chemother. Pharmacol.*, 52(6):435–441, Dec 2003.

[26] J. C. Panetta, M. N. Kirstein, A. J. Gajjar, G. Nair, M. Fouladi, and C. F. Stewart. A mechanistic mathematical model of temozolomide myelosuppression in children with high-grade gliomas. *Math Biosci*, 186(1):29–41, Nov 2003.

[27] 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.

[28] A. S. Zandvliet, J. H. Schellens, J. H. Beijnen, and A. D. Huitema. Population pharmacokinetics and pharmacodynamics for treatment optimization in clinical oncology. *Clin Pharmacokinet*, 47(8):487–513, 2008.

Electronic Supplemental Material for "Personalized oncology with artificial intelligence: The case of temozolomide"

Nicolas Houy* François Le Grand†

July 8, 2019

1 PK/PD Model

The equations for the PK/PD model we use are exactly the same as in [1] and can be found in the supporting information at `https://doi.org/10.1371/journal.pone.0199076.s001`.

The parameters values are also the same but for the volume of distribution V that consider with its dependence to BSA as in [2]. The volume of distribution V is a random variable that statistically[1] depends on BSA as follows

$$V = \exp\left(a_{BSA}.BSA + b_{BSA} + \epsilon\right).$$

These original parameters' values are

- BSA is a normal random variable with mean 1.19 and standard deviation 0.4,
- ϵ is a normal random variable with mean 0 and standard deviation 0.145,
- $a_{BSA} = 1.09$,
- $a_{BSA} = 1.34$.

*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]The dependence is statistical and then observing BSA is informative and allows to have a narrower belief about V but it is not fully informative.

Moreover, whenever a patient is treated, a noisy signal about parameters k_a, k_e and V is fed to the algorithm that updates the subjective beliefs about the unknown parameters k_a, k_e and V. The noise with which the information is given corresponds to the inter-occasion variability given in [2]:

- a signal for $\log(k_a)$ is drawn from a centered normal distribution with standard deviation 0.97,
- a signal for $\log(k_e)$ is drawn from a centered normal distribution with standard deviation 0.3617,
- a signal for $\log(V)$ is drawn from a centered normal distribution with standard deviation 0.3662.

2 The OPP algorithm

The OPP algorithm is similar to the H algorithm given and extensively commented in [1]. The only differences lie in the following points.

- At the beginning, of the algorithm, we have subjective distribution functions for k_a, k_e and V that equals the objective ones. We observe BSA (which is formally a noisy observation of V because of ϵ) and update (with Bayes rule) our belief about V.
- Whenever a population is generated for policy simulation, it is generated with k_a, k_e and V drawn from the current subjective distribution functions.
- Whenever temozolomide is administered, we have a noisy (as described in the previous section) observation of k_a, k_e and V. Notice that because of this observation, temozolomide administration has an informative value[2] that is taken into account by our algorithm.

3 Computing the difference between protocols

Our in-silico approach allows us to administer different protocols to the same (fictive) population – or more precisely to two fictive populations having exactly the same characteristics. From a statistical point of view, this enables us to use paired tests, whose power increases with the size of the population.

[2]Loosely speaking, in our model, temozolomide administration has the virtue to cure but also to gain information that allows more precise beliefs of the parameters and hence a better optimization for the future. The latter effect is the value of information.

To show it, consider a variable x (such as tumor size at terminal date for instance). We assume that two identical populations of size n to which two protocols A and B have been administered. We denote $(x_1^A, \ldots, x_n^A)$ and $(x_1^B, \ldots, x_n^B)$ the distribution of x for the two populations, respectively. The respective averages are defined as:

$$\overline{x}^A = \frac{1}{n}\sum_{i=1}^{n} x_i^A \quad \text{and} \quad \overline{x}^B = \frac{1}{n}\sum_{i=1}^{n} x_i^B$$

We assume that the variable x is iid in the population (in particular the outcome for one patient does affect the one for another one), with $\text{var}(x_i^A) = \sigma_A^2$, $\text{var}(x_i^B) = \sigma_B^2$ (for all i), and $\text{cov}(x_i^A, x_j^B) = \rho_{A,B}\sigma_A\sigma_B 1_{i=j}$ ($1_{i=j} = 1$ if $i = j$ and 0 otherwise). We deduce that the variance of the difference $\overline{x}^A - \overline{x}^B$ is equal to:

$$\begin{aligned}
\text{var}(\overline{x}^A - \overline{x}^B) &= \text{var}(\overline{x}^A) + \text{var}(\overline{x}^B) - 2\text{cov}(\overline{x}^A, \overline{x}^B) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \underbrace{\text{var}(x_i^A)}_{=\sigma_A^2} + \frac{1}{n^2}\sum_{i=1}^{n} \underbrace{\text{var}(x_i^B)}_{=\sigma_B^2} - 2\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \underbrace{\text{cov}(x_i^A, x_i^B)}_{=\rho_{A,B}\sigma_A\sigma_B 1_{i=j}} \\
&= \frac{\sigma_A^2 + \sigma_B^2 - 2\rho_{A,B}\sigma_A\sigma_B}{n},
\end{aligned}$$

which is decreasing with n. In other words, the larger the population size, the more significant the difference between the two populations. Indeed, the Z-test for assessing how the two means differ from each other, can be expressed as $(\overline{x}^A - \overline{x}^B)/\sqrt{\text{var}(\overline{x}^A - \overline{x}^B)}$ (which under the null hypothesis that the two means are identical, can be assumed to follow a normal distribution given that n is quite large in our paper).

References

[1] N. Houy and F. Le Grand. Optimal dynamic regimens with artificial intelligence: The case of temozolomide. *PLoS ONE*, 13(6):e0199076, 2018.

[2] J. C. Panetta, M. N. Kirstein, A. Gajjar, G. Nair, M. Fouladi, R. L. Heideman, M. Wilkinson, and C. F. Stewart. Population pharmacokinetics of temozolomide and metabolites in infants and children with primary central nervous system tumors. *Cancer Chemother. Pharmacol.*, 52(6):435–441, Dec 2003.