The value of monitoring to control evolving populations

Andrej Fischer[a], Ignacio Vázquez-García[a,b], and Ville Mustonen[a,1]

[a]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom; and [b]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, United Kingdom

Populations can evolve to adapt to external changes. The capacity to evolve and adapt makes successful treatment of infectious diseases and cancer difficult. Indeed, therapy resistance has become a key challenge for global health. Therefore, ideas of how to control evolving populations to overcome this threat are valuable. Here we use the mathematical concepts of stochastic optimal control to study what is needed to control evolving populations. Following established routes to calculate control strategies, we first study how a polymorphism can be maintained in a finite population by adaptively tuning selection. We then introduce a minimal model of drug resistance in a stochastically evolving cancer cell population and compute adaptive therapies. When decisions are in this manner based on monitoring the response of the tumor, this can outperform established therapy paradigms. For both case studies, we demonstrate the importance of high-resolution monitoring of the target population to achieve a given control objective, thus quantifying the intuition that to control, one must monitor.

stochastic optimal control | adaptive cancer therapy | drug resistance

The progression of cancer is an evolutionary process of cells driven by genetic alterations and selective forces (1). The frequent failure of cancer therapies, despite a host of new targeted cancer drugs, is largely caused by the emergence of drug resistance (2). Cancer therapy faces a real dilemma: the more effective a new treatment is at killing cancerous cells, the more selective pressure it provides for those cells resistant to the drug to take over the cancer population in a process called competitive release (3, 4).

A genetic innovation conferring resistance can either be already present as standing variation or in close evolutionary reach, via de novo mutations. The probability of these events is often proportional to the genetic diversity of the tumor. Therefore, resistance is a problem especially for genetically heterogeneous cancers (5). This diversity can be the result of a variable microenvironment, with different pockets of acidity, blood supply, and geometrical constraints of surrounding tissue (2). Also, late-stage cancers not only carry the cumulative archaeological record of their evolutionary history (6) but can also become genetically unstable and fall victim to chromothripsis (7), kataegis (8), and other disruptive mutational processes (9, 10). Thus, the probability of treatment success is higher in genetically homogeneous and/or early-stage cancers (11). Taken together, these considerations place emphasis on early detection of tumors.

In cases where early detection is not achieved, the pertinent question is how to avoid treatment failure in the presence of genetic heterogeneity, which seems to be the norm for most solid cancers. One obvious attempt is to make treatments more complex and thus put the resistance mechanisms out of reach of the tumor. In combination therapy, the tumor is simultaneously treated with two or more drugs that would require different, possibly mutually exclusive, escape mechanisms for cells to become resistant. This approach has proven to be successful in the treatment of HIV and is discussed as a possible model also for cancer (12). In the context of cancer, this form of personalized therapy is not yet widely realized, mainly because of the much richer repertoire of genetic variation and adaptability of cancer cells and a comparable shortage of drugs targeting distinct

biological pathways. For a recent study of the conditions under which combination therapy is expected to be successful in cancer, see ref. 13.

For application of single drugs, there are a number of studies that concentrate on how the therapeutic protocol itself can be optimized. It was realized that all-out maximum tolerated dose chemotherapy is not the only, or necessarily the best, treatment strategy (14). Alternative dosing schedules have been proposed such as drug holidays, metronome therapy (15), and adaptive therapy (16). The realization of Gatenby et al. in ref. 16 is that cancer, as a dynamic evolutionary process, can be better controlled by dynamically changing the therapy, depending on the response of the tumor. Their protocol of reducing the dose while the tumor shrinks and increasing it under tumor growth showed a drastic improvement of life expectancy in mice models of ovarian cancer (16). Furthermore, Gatenby et al. made the important conceptual step of reformulating cancer therapy to be not necessarily about tumor eradication. Instead, dynamic maintenance of a stable tumor size can also be a preferable outcome.

Motivated by this experiment, we conjecture that there are substantial therapy gains in optimal applications of existing drugs, as of yet underexploited. As a first step toward using this potential we would like to formalize the intuition of Gatenby et al. To this extent, we aim to establish a theoretical framework for the adaptive control of evolving populations. In particular, we connect the idea of adaptive therapy to the paradigm of stochastic optimal control, also known as Markov decision problems. For other applications of stochastic control in the context of evolution by natural selection see refs. 17, 18. A stochastic treatment is necessary due to the nature of evolutionary dynamics where fluctuations (so-called genetic drift) matter even in large populations. For instance, the dynamics of a new beneficial mutation is initially dominated by genetic drift before it becomes established (19). Stochastic control is a well-established field of research which provides not only a natural language for framing the task of cancer therapy, but also a set of general purpose

Significance

Evolution of drug resistance, as observed in bacteria, viruses, parasites, and cancer, is a key challenge for global health. We approach the problem using the mathematical concepts of stochastic optimal control to study what is needed to control evolving populations. We focus on the detrimental effect of imperfect information and the loss of control it entails, thus quantifying the intuition that to control, one must monitor. We apply these concepts to cancer therapy to derive protocols where decisions are based on monitoring the response of the tumor, which can outperform established therapy paradigms.

PHYSICS

EVOLUTION

techniques to compute an optimal control or therapy regimen for a given dynamical system and a given control objective. Although we demonstrate the main steps in this program, we focus on the detrimental effect of imperfect information and the loss of control it entails, thus quantifying the intuition that to control, one must monitor. The informational value of continued monitoring is a natural concept for controlled stochastic systems, whereas in deterministic models successful control usually does not rely on sustained observations.

We first introduce the concepts of stochastic optimal control using a minimal evolutionary example: how to keep a finite population polymorphic under Wright–Fisher evolution by influencing the selective difference between two alleles. If perfect information about the population is available, the polymorphism can be maintained for a very long time. We will show how imperfect information due to finite monitoring can lead to a quick loss of control and how some of it can be partially reclaimed by informed preemptive control strategies. We then move to our main problem and introduce a minimal stochastic model of drug resistance in cancer that incorporates features such as variable population size, drug-sensitive and -resistant cells, a carrying capacity, mutation, selection, and genetic drift. After computing the optimal control strategies for a few important settings under perfect information, we demonstrate the effect of imperfect monitoring. If only the total tumor size can be monitored, we show how a control strategy emerges that can adaptively infer, and thus exploit, the inner tumor composition of susceptible and resistant cells.

Controlling Evolving Populations

One can think about cancer therapy as the attempt to control an evolving population by means of drug treatment. Typically, the drug changes some of the parameters of the evolutionary process, such as the death rate of drug-sensitive cells. With application of the drug, one can thus actively influence the dynamics of the stochastic process and change its direction. All this happens with a concrete aim, such as to minimize the total tumor burden in the long term. To introduce some of the concepts of stochastic optimal control, we use an example with a nontrivial control task.

Imagine a biallelic and initially polymorphic population of constant size N under the Wright–Fisher model of evolution (20), i.e. binomial resampling of the population in each generation. The A allele confers a selective advantage of size $\sigma = N(f_A - f_B) \gg 1$ over the B allele and will, without intervention, eventually take over the entire population (f denotes reproductive rates; see Fig. 1A). Assuming mutation to be negligible, the task at hand is to avoid, or at least delay, such a loss in diversity. Now assume that we can change the selection coefficient externally by a quantity $u \in [-u_c, 0]$ in the form $\sigma \rightarrow \sigma + u$. Ideally, we would have $|u_c| > \sigma$, but this is not a necessary condition. In this setting, the control problem can be stated as follows: for a population with an initial A-allele frequency of $x_0 = n_0/N \in (0,1)$, what is the optimal control protocol $\overline{u}_{0:T}(x_0)$ that maximizes the probability that the polymorphism is still alive after a time of T generations?

$$\overline{u}_{0:T}(x_0) = \underset{u_{0:T}}{\mathrm{argmax}}\, P(x_T \in (0,1)|x_0), \qquad [1]$$

where the maximization is to be done over all possible sequences of controls $u_{0:T} = u_0 u_1 \ldots u_T$ ($u_t \in [-u_c, 0]$). Note that the stochastic nature of the process makes the control optimal only in the sense of expected outcomes. Individual realizations might well fall short of, or exceed, the implied mean survival time. It is helpful to picture a large ensemble of such populations, all starting off at frequency x_0, all being individually nudged by selection according to a (yet to be found) optimal protocol.
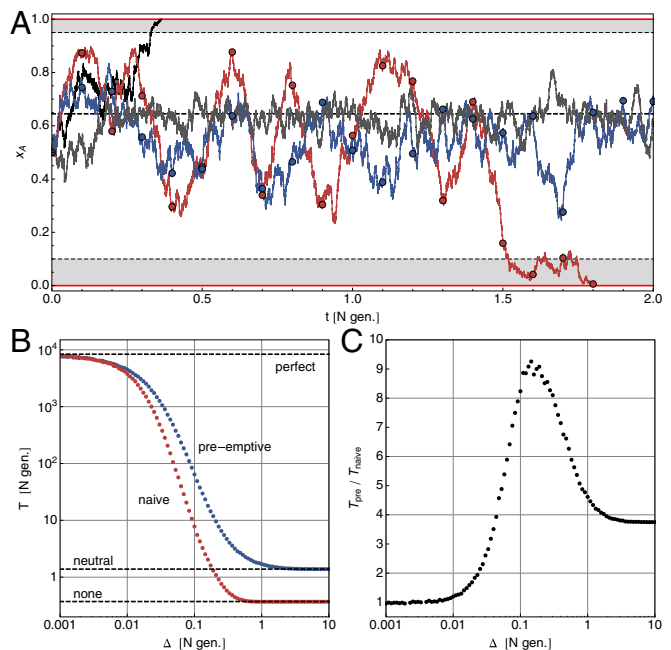
Fig. 1. Optimal control of a finite population under Wright–Fisher evolution to maintain a polymorphism. The intrinsic selection coefficient is $\sigma = 10$ and control shifts selection to $\sigma + u$. (A) Sample trajectories starting at $x_0 = 0.5$: without control ($u \equiv 0$, black line) the polymorphism is lost on a time scale of $1/\sigma$. With optimal control under perfect information (gray line, $\sigma + u = \sigma = 10$ for $x < x_c \approx 0.64$, else $\sigma + u = \sigma - u_c = -20$), it can be maintained for an average of 8,000 N generations. With finite monitoring ($\Delta = 0.1$, measurements x_i at circles), naive control ($u \equiv \overline{u}(x_i)$), red line, is prone to overshooting, whereas preemptive control (blue line) tries to avoid this by switching to a neutral regime after a short time. (B) Loss of control under finite monitoring: as Δ grows, so does the probability that the polymorphism is already lost at the next measurement. Shown is the mean survival time over 5,000 trajectories with $N = 10^4$ on a logarithmic scale. (C) Under preemptive control, some of the loss of control can be regained, especially for intermediate values of Δ.

When a trajectory hits one of the boundaries before the final time T, it is lost. The optimal control can thus also be seen to minimize this cost of extinction. The standard technique to solve problems of this kind is to use a dynamic programming ansatz. Assuming that a partial problem starting from some intermediate point (x_t, t) ($0 < t < T$) has already been solved, we define the cost-to-go $J(x_t, t)$ as the corresponding optimal cost. From this definition follows a backward-recurrence relation for J: the cost-to-go at (x_t, t) is the expected cost-to-go one step later, where the average depends via the propagator W on the control decision u_t at the current time (21):

$$J(x_t, t) = \min_{u_t} \sum_{x'} J(x', t+1) W(x'|x_t; u_t). \qquad [2]$$

The absorbing boundary conditions take the form $J(0, t) = J(1, t) = T - t$. The right-hand side of Eq. **2** can also include a term $V(x_t, u_t)$ that describes the potential cost to be at x and the control cost to apply u. Here, both are assumed zero and cost is paid only at the boundaries. For Wright–Fisher evolution, the transition matrix W can be expressed as the probability under binomial sampling to draw $n'_A = N x'$ individuals of the A allele. The crucial computational advantage of this relation is that the hard optimization in the space of all control protocols $u_{0:T}$ is exchanged for a simple scalar optimization over u_t. In statistical physics, this technique is referred to as the transfer-matrix method. The intuitive interpretation of Eq. **2**

is that the decision for a control now relies on future controls to be carried out optimally. In practice, the results of the local optimizations $(\overline{u}(x_t, t) = \underset{u_t}{\operatorname{argmin}} \dots)$ constitute the optimal control to be applied when the system is at x_t at time t. In many applications (21), it is also useful to consider a receding time horizon, such that $\overline{u}(x, 0)$ is a stationary control. In general, the solution to Eq. **1** is not guaranteed to be unique. However, even degenerate optimal controls would achieve the same optimal value of the control objective (22).

Throughout this study, we apply the diffusion approximation ($N \to \infty$, while σ, u_c fixed and $\tau = t/N$), where the infinitesimal form of Eq. **2**, the so-called Hamilton–Jacobi–Bellman equation (21, 23), is valid.

$$-\partial_\tau J(x, \tau) = \min_{u_\tau} x(1-x) \left[(\sigma + u_\tau)\partial_x + \frac{1}{2}\partial_x^2 \right] J(x, \tau). \quad [3]$$

We propose an ansatz to solve the control problem for the Wright–Fisher example, and confirm it by direct numerical application of Eq. **2** (see also Fig. S1). For a treatment of finite mutation rate, see *SI Text* and Fig. S2.

Optimal Control of a Wright–Fisher Population with Perfect Monitoring. The optimal control function $\overline{u}(x_t, t)$ maximizes the probability that a polymorphism is still present after a time T. In the infinite horizon time limit $T \to \infty$, where the optimal control becomes stationary $\overline{u}(x)$, we expect it to also maximize the mean first passage time $\langle T_x^{(0,1)} \rangle$ out of the interval $(0, 1)$ (for any fixed starting position x). In Eq. **3**, the control variable u appears only linearly. This means that only the two extreme control strengths are ever used to steer the system. This particular type of control is called bang-bang. It follows that the control profile $\overline{u}(x)$ will have the form of a step function with critical frequency x_c,

$$\sigma + \overline{u} = \begin{cases} \sigma, & x < x_c \\ \sigma - u_c, & x \geq x_c \end{cases}. \quad [4]$$

The only remaining parameter is the critical threshold $x_c(\sigma, u_c)$. To find an expression for the objective function $\langle T^{(0,1)} \rangle(x_c, \sigma, u_c)$, we can consider the optimally controlled system as the simplest example of evolution under frequency-dependent, piecewise constant selection $\sigma + \overline{u}(x)$. The mean first passage time can then be found analytically using standard methods for stochastic processes (24) (*SI Text* and Fig. S1). At the correct threshold and with strong selective forces ($\sigma \gg 1$ and $\sigma - u_c \ll -1$), the gains are substantial and the polymorphism can be maintained for very long times.

Loss of Control due to Imperfect Monitoring. The main assumption made so far was that perfect information is available about the state of the system in the form of continuous (in time), synchronous (without delay), and exact (without error) measurements of x. These requirements are impossible to achieve in practice, when monitoring is always imperfect. As we will see, when the assumption of perfect information is relaxed, not only is control over the system lost, but the control profile $\overline{u}(x)$ also ceases to be optimal. Rather than turning to the theory of partially observable Markov decision problems (25), we will use numerical analysis to demonstrate the effect of monitoring with finite resolution in time (relaxing the first condition).

Consider the situation where measurements of the frequency x are given only at discrete times $\{\tau_i\}$, whereas no information is available during the intervals of length $\Delta = \tau_{i+1} - \tau_i$. The immediate question is, given a measurement x_i, what control should one apply while waiting for the next measurement? The perfect-information control $\overline{u}(x_i)$ is correct only initially, and thus only in the limit $\Delta \to 0$. However, it is intuitively clear that a naive

protocol, applying $\overline{u}(x_i)$ during the entire interval Δ, cannot be optimal, because it does not anticipate the dynamics of x under this regime (see the decrease in survival time in Fig. 1B). For example, for $0 < x_i < x_c$ the initial control is $u = 0$, $\sigma + u > 0$ and the frequency will, on average, increase and eventually cross the threshold x_c. If one could observe the population at that point, the control should be switched to $u = -u_c$ until x crosses x_c from above. The total result of the naive strategy is to amplify fluctuations due to this overshooting.

Playing-To-Win vs. Playing-Not-To-Lose. Without a continuous flow of observations as input, a preemptive control protocol $u^*(\tau, x_i)$ during the interval Δ must be precomputed and then faithfully carried out. In the discrete-time (Wright–Fisher evolution) setting, there are $N\Delta$ generation updates until the next measurement and therefore $2^{N\Delta}$ different bang-bang protocols to choose from. However, the example above suggests to search for a preemptive control in a much smaller space, namely within those protocols that start with either $u^* = 0$ or $u^* = -u_c$ and then switch, at some later time $\tau_c(x_i, \Delta)$, to a neutral regime with $u^* = -\sigma$ (if available). The size of this space is only $2 N\Delta$ (we note that for finding an optimal control one would need to search protocols in the full space). The effect of such a control scheme is to move the population to a safe place and then try to keep it there. There are two important observations: first, this informed control outperforms the naive protocol significantly, especially for intermediate values of Δ (Fig. 1 B and C); second, the safe parking position moves away from the boundary toward $x = 0.5$ for larger values of Δ (Fig. S3). This shift from an aggressive control strategy under perfect information ($\Delta = 0$, x_c close to a boundary) to a more and more conservative one (aiming for $x = 0.5$ and trying to stay there) can be summarized as playing-to-win vs. playing-not-to-lose.

A similar loss of control can be expected for other types of monitoring imperfections and is a general feature of stochastic optimal control. It is important to note that the perfect-information control problem, and its solution $\overline{u}$, is a typical starting point for the analysis. The naive control protocol above is indeed optimal for $\Delta \to 0$, and still a very good option for $\Delta \ll 1$ (in units of N generations). In most cases, as we will see in the adaptive cancer therapy model below, finding $\overline{u}$ is challenging in itself and can be a good guidance for finding well-performing control protocols even under imperfect conditions.

Application to Adaptive Cancer Therapy

Here we first introduce a minimal stochastic model of drug resistance in cancer. For different qualitative regimes, we then find the optimal adaptive therapy with perfect information. Finally, we extend these ideas to the case where only the total cell population size can be observed but no readout of the fractions of susceptible and resistant cells is available. In the context of models of the cell cycle, deterministic control theory has previously been applied to find optimal cancer treatment protocols (e.g. refs. 26–28). However, for the key concepts of this study—adaptive therapy and finite monitoring—stochastic control theory is needed and in fact leads to control protocols that exploit fluctuations.

A Minimal Model of Drug Resistance in Cancer. The desired features of a minimal model of drug resistance in cancer include: (*i*) a variable tumor cell population size N, (*ii*) at least two cell types, drug-sensitive and drug-resistant, (*iii*) a carrying capacity K that describes a (temporary) state of tumor homeostasis, and (*iv*) the possibility for mutation and selection between cell types. Control over the tumor can be applied via a drug that changes the evolutionary dynamics by increasing, for example, the death rate of sensitive cells. We will assume here, as others have done in the context of cancer (11), a well-mixed cell population where the

PHYSICS

EVOLUTION

birth (or rather duplication) rate of cells is regulated by the carrying capacity. The dynamics of the model we have chosen here is encapsulated in the following birth and death rates for sensitive and resistant cells:

$$B_i(n_s, n_r) = \frac{(1 + g_i)n_i}{1 + g\frac{N}{K} + a\frac{n_s}{K}} + \mu_0(n_{\bar{i}} - n_i),$$

$$D_i(n_s, n_r) = n_i(1 + F_i(u)), \quad i \in \{s, r\}, \quad \bar{s} = r, \quad \bar{r} = s,$$ [5]

with free growth rates $g_s = g + a$, $g_r = g$, growth advantage $a \geq 0$, and mutation rate μ_0 between cell types. F_i encodes the effect of the drug ($u = 1$) or its absence ($u = 0$) on cell type i. For $N \ll K$, the absolute growth rates are $g_i - F_i$. A drug effect of the form

$$F_s(u) = u f_s \quad \text{and} \quad F_r(u) = (1 - u)f_r,$$ [6]

renders the drug effective if $f_s > g + a$. The value $f_r = 0$ corresponds to drug resistance as such, but $f_r > g$ implies that resistant cells thrive under the drug and are drug-addicted. Such an effect has been observed in mice with melanoma carrying oncogene *BRAF* mutation treated with vemurafenib (29). Altogether, sensitive and resistant cells initially grow exponentially until the total population size $N = n_s + n_r \approx K$. For some of the protocols that we derive, it is important that the tumor is close to or can reach carrying capacity during treatment (e.g. patients who are inoperable or whose primary tumors are not resectable). At that stage, competition for resources, space, etc. becomes fierce. If sensitive cells have a differential growth advantage $a > 0$ (they might not have to maintain an expensive resistance mechanism), resistant cells will eventually be removed from the tumor or reduced to a small fraction (of size μ_0/a). In reality, this scenario might not materialize, as the next mutation could propel the tumor into a new phase of exponential growth.

For the stochastic version of this process we can assume independent and individual birth and death events with the above probabilities per unit time. In analogy to the Wright–Fisher binomial update rule, here we can use a Poisson-like update.

$$n_i \rightarrow n_i' = n_i + \Delta n_i, \quad \text{with } \Delta n_i = \Delta n_i^+ - \Delta n_i^-$$ [7]

$$\Delta n_i^+ \sim \text{Pois}(B_i), \quad \Delta n_i^- \sim \text{Pois}(D_i), \quad \langle \Delta n_i \rangle = B_i - D_i.$$

The total increment Δn_i follows a Skellam distribution. When we let $K \rightarrow \infty$ while fixing the combinations $\gamma \equiv Kg$, $\alpha \equiv Ka$, $\mu \equiv K\mu_0$, and $\phi_i = Kf_i$ and setting $t = \tau/K$, the system is described by a Fokker–Planck evolution equation (30) for the distribution $P(x_s, x_r, t)$ with $x_i \equiv n_i/K$ (*SI Text*). This scaling exercise is mainly important because it allows to relate systems with small K (hundreds to thousands, as required for numerical analysis) to systems with large K ($\gtrsim 10^8$, as present in real cancers).

Optimal Cancer Therapy with Perfect Monitoring. With the minimal model of drug resistance in cancer introduced above, we can start the program of stochastic optimal control to compute adaptive therapy protocols. The first task is to define the goal of such a program: what is the quantity one aims to optimize? One very important objective is to maximize the (expected) time until the cancer proceeds to the next, possibly lethal stage. This could mean the emergence of a new cell type with a much higher carrying capacity, e.g. with metastatic potential. We will denote this critical event simply with a "driver" event or "metastasis." The rate of metastasis emergence is a combination of tumor size and the rate ν_0 (per cell and generation) for the necessary features to appear via mutation.

Earlier, the optimal control for the Wright–Fisher evolution example turned out to be a piecewise constant function of allele

frequency. Here, we need to find a control profile $\bar{u}(n_s, n_r)$. With perfect information, we would know n_s and n_r at all times and would base the control decision adaptively on these measurements. In analogy to Eq. **1**, the control objective can be expressed as

$$\bar{u}_{0:T}(n_{s0}, n_{r0}) = \underset{u_{0:T}}{\text{argmax}} \exp\left(-\nu_0 \sum_{t=0}^{T}(n_{st} + n_{rt})\right),$$ [8]

where the right-hand side is the probability that metastasis has not yet happened after T generations (for $\nu_0 K \ll 1$). Here, the control objective is a nonlinear function of the entire trajectory $(n_s, n_r)_{0:T}$. As such, it is the simplest manifestation of a so-called risk-sensitive control problem (see e.g. ref. 31). The above formulation assumes a finite (receding) horizon time T and also that control itself is cost-free. In cancer therapy, especially chemotherapy, this is certainly not the case: the side effects of treatment incur a considerable cost in terms of life quality and medical care. The difficulty, however, lies in quantifying these control costs in a manner that would make them comparable to the potential costs considered here. A simple extension of the above control objective to include such a control cost is demonstrated in *SI Text* and Fig. S4. The recurrence equation for the cost-to-go $J(n_s, n_r, t)$ for the control objective above is given by (31)

$$J(n_s, n_r, t) = e^{-\nu_0(n_s + n_r)} \max_{u \in \{0,1\}} \langle J(t+1; u) \rangle,$$ [9]

$$\langle J(t'; u) \rangle \equiv \sum_{n_s', n_r'} J(n_s', n_r', t') \, W(n_s', n_r' | n_s, n_r; u),$$ [10]

with boundary condition $J(T) = 1$. The (microscopic) transition matrix W is the product of the two Skellam distributions resulting from Eq. **7** (including boundary conditions). With this equation, we can solve the dynamic programming task numerically for moderate values of K (see Fig. S5). For the numerical analysis, we have to introduce an upper bound $\tilde{N} \gg K + \sqrt{K}$ for the population size. The resulting control profiles for a number of different parameter regimes is shown in Fig. 2.

In the case of $\phi_r = 0$, resistant cells are unaffected by the drug. If maintenance of the resistance mechanism is costly ($\alpha > 0$), the only way that they can be removed from the population is when selection can act against them. This only happens at $N \sim K$ and with $u = 0$ (no drug). If this can take place before the next driver typically appears (if $\alpha = Ka \gg KN\nu_0 \sim K^2\nu_0 \equiv \nu$), then the optimal control protocol is to postpone treatment until the resistant cells are sufficiently cleared from the system so that genetic drift has a chance of removing the remaining part (Fig. 2*A*). However, this parameter regime of very high selection against resistance and/or very low rate of driver mutation, and therefore this therapy option, is not realistic for cancer. For higher values of ν/α, the optimal strategy is to apply the drug earlier (Fig. 2*B*). This procedure can lead to cycles of tumor size reduction followed by regrowth, with the overall effect of extending the time until metastasis.

If $\phi_r > \gamma$ (and $\phi_s > \gamma + \alpha$), resistant cells are actually drug-addicted and thrive only in its presence. Such a situation would be easy to control with perfect information about n_s and n_r. For example, if mutation between cell types is rapid ($\mu \gg 1$), a majority rule is optimal (Fig. 2*C*). For a lower mutation rate, the optimal protocol first tries to amplify one cell type before switching to an environment that is deadly for most cells (Fig. 2*D*).

The effectiveness of different therapy protocols is compared in Fig. 3 with 1,000 stochastic forward simulations (with $K = 10^4$) for the parameter setting of Fig. 2*D*. Whereas no therapy ($u = 0$) and all-out therapy ($u = 1$) both ultimately end with the occurrence of metastasis, adaptive therapy can bring the tumor
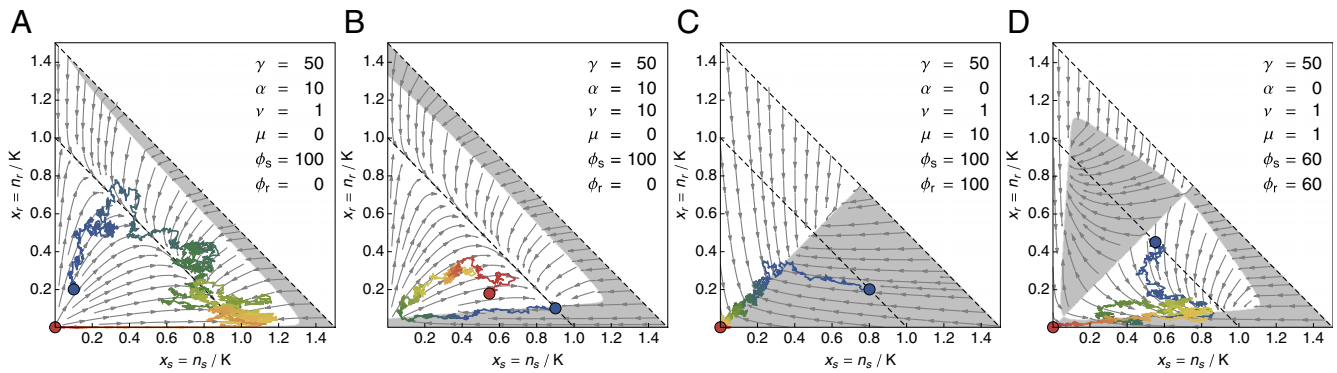
Fig. 2. Control of a tumor cell population. (*A–D*) The optimal control profile under perfect information about n_s and n_r for different parameters of the cancer model. In the white areas, $\bar{u} = 0$ (no drug), whereas in the gray areas, $\bar{u} = 1$ (with drug). The arrows indicate the deterministic flow. All profiles were calculated via Eq. 9 with $T = K/\nu$ generations and $K = 500$ with an absorbing boundary at $N = 750$. The sample trajectories were simulated with $K = 10^4$ and controlled according to these profiles. The coloring of the trajectories shows the temporal evolution from blue to red. (*A*) When selection against resistance is stronger than driver emergence, $\alpha \gg \nu$, the optimal protocol is to wait until resistant cells are cleared from the system before the drug is applied. (*B*) For higher driver emergence rates, the drug is applied earlier, which can lead to cycles. (*C*) For drug-sensitive ($\phi_s \gg \gamma$) and drug-addicted cells ($\phi_r \gg \gamma$) with high mutation ($\mu \gg 1$), the control in the symmetric case ($\phi_s = \phi_r$) is a simple majority rule and very effective. (*D*) For smaller mutation ($\mu = 1$), the optimal strategy first homogenizes the tumor before trying to remove it.

size down to zero in the majority of cases. In metronome therapy, the drug is applied (withheld) for fixed time intervals τ_{on} (τ_{off}). With numerically optimized values for the time intervals, metronome therapy is quite competitive compared with adaptive therapy.

All these control strategies require perfect information, not only in the sense of the earlier Wright–Fisher example (continuous, synchronous, and exact), but also in terms of the inner tumor composition $N = n_s + n_r$, which presupposes that sensitive and resistant cells can be distinguished.

Loss of Therapy Efficacy due to Low-Resolution Monitoring. There are very few cases where the genetic basis for a drug-resistance mechanism is known and can be specifically monitored (29, 32). In most cases the regrowth of the tumor under the drug is observed without understanding the exact biological processes responsible for the resistance. Here we aim to find rational control strategies when only the total tumor cell population size can be monitored. The adaptive therapy protocol that was applied by Gatenby et al. in ref. 16 (coupling the drug concentration to the tumor size) is one example of such a strategy.

Consider the situation where only the total population size $N = n_s + n_r$ can be (perfectly) monitored, whereas the dynamical laws in Eqs. **5–7** and all parameter values are known. Under these circumstances, the perfect-information optimal control profiles from the last section cannot be used directly. However, there is still valuable information available. The response $N_\tau \to N_{\tau+\Delta\tau}$ of the tumor size to a control choice over a time interval $\Delta\tau$ can give an indication of the inner tumor composition. As we have seen earlier, the length of time interval $\Delta\tau$ should be shorter than all other intrinsic time scales to enable control. The procedure we follow is to first derive an effective propagator $W_{eff}(N_{\tau+\Delta\tau}|N_{\tau-\Delta\tau}, u_{\tau-\Delta\tau}; u_\tau)$ (*SI Text*) and then repeat the cost-to-go calculation of Eqs. 9 and 10. This propagator takes into account not only the current size N_τ, but also the last measurement $N_{\tau-\Delta\tau}$ and the last control decision $u_{\tau-\Delta\tau}$. It follows from the microscopic W used in Eq. 10 by integrating over the internal degrees of freedom n_s and n_r at the three time points. Accordingly, the control profile is now a function of $(N_{\tau-\Delta\tau}, u_{\tau-\Delta\tau}, N_\tau)$. For the parameter values leading to the majority rule in Fig. 2C, the new control profile is shown in Fig. S6. The drug regimen ($u = 0$ or $u = 1$) is maintained as long as the tumor size decreases sufficiently. At the first sign of possible reversal, the regimen is switched.

Discussion

We used stochastic control theory to quantify optimal control strategies for models of evolving populations. We hope our results will lead to interesting new designs of microbial and cancer cell evolution experiments where feedback plays a central role in achieving a given control task. We further demonstrated how control can be maintained with finite resources, when the monitoring necessary for adaptive control is imperfect. These strategies all depend on our ability to anticipate evolution, i.e. on a knowledge of the relevant equations of motion and their parameter values. For cancer, such detailed knowledge of evolutionary dynamics is certainly not yet available. Sequencing technologies are facing up to the challenge of tumor control with finite information, already accelerating progress in the monitoring of serial biopsies of tumors, circulating tumor cells, or cell-free tumor DNA in the bloodstream (33, 34). Once such time-resolved data become prevalent, we can start to learn and improve dynamical tumor models and compute their optimal control
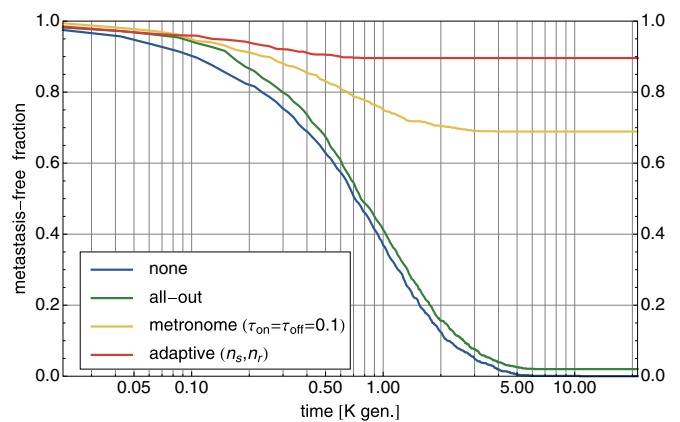
Fig. 3. Comparison of cancer therapies. For the parameter setting of Fig. 2*D*, different therapies are compared via 1,000 forward simulations with $K = 10^4$. Shown is the fraction of runs that have not yet developed a metastasis mutation as a function of elapsed time. All-out maximum dosage therapy ($u \equiv 1$) is only slightly better than no therapy ($u \equiv 0$) in avoiding metastasis. Much better is metronome therapy with $\tau_{on} = \tau_{off} = 0.1$ K generations, with almost 70% success rate (values numerically optimized). Adaptive therapy removes the tumor in close to 90% of runs.

strategies. For instance, genetic heterogeneity within the tumor is now becoming quantifiable from sequencing data via computational inference (35–37). Heterogeneity and subclonal dynamics have been found to have an impact on treatment strategy selection (38). Furthermore, all other available sources of clinical data, such as medical imaging, can provide additional information on cellular phenotypes and should be integrated into personalized and data-driven tumor control (see e.g. ref. 39 for imaging data-based computational modeling of pancreatic cancer growth dynamics to guide treatment choice and ref. 40 for integrative analysis of imaging and genetic data).

Beyond cancer, the need to control evolving populations is a key global health challenge as resistant strains of bacteria, viruses, and parasites are spreading (41, 42). Any long-term success in controlling evolution depends, at the very least, on mastering the following components. Firstly, on a quantitative understanding of the underlying evolutionary dynamics. Progress in the understanding is best demonstrated by predicting evolution; this has so far proven difficult, even in the short term.

Nevertheless, new population genetic approaches applied to data are promising—see influenza strain prediction in ref. 43. Secondly, the success of control will depend on the availability of a sufficient arsenal of non–cross-resistant therapeutic agents. These therapeutics should be combined with the ability to decide an appropriate drug regimen given the genetic and phenotypic structure of the population. Large-scale drug vs. cell line screens are systematically pushing this component forward (see e.g. ref. 44). And finally, long-term success depends on the ability to monitor the evolution of a target population and act rationally based on this information, the topic of this paper.

1. Stratton MR, Campbell PJ, Futreal PA (2009) The cancer genome. *Nature* 458(7239): 719–724.
2. Gillies RJ, Verduzco D, Gatenby RA (2012) Evolutionary dynamics of carcinogenesis and why targeted therapy does not work. *Nat Rev Cancer* 12(7):487–493.
3. Wargo AR, Huijben S, de Roode JC, Shepherd J, Read AF (2007) Competitive release and facilitation of drug-resistant parasites after therapeutic chemotherapy in a rodent malaria model. *Proc Natl Acad Sci USA* 104(50):19914–19919.
4. Greaves M, Maley CC (2012) Clonal evolution in cancer. *Nature* 481(7381):306–313.
5. Gerlinger M, Swanton C (2010) How Darwinian models inform therapeutic failure initiated by clonal heterogeneity in cancer medicine. *Br J Cancer* 103(8):1139–1143.
6. Nik-Zainal S, et al.; Breast Cancer Working Group of the International Cancer Genome Consortium (2012) The life history of 21 breast cancers. *Cell* 149(5):994–1007.
7. Forment JV, Kaidi A, Jackson SP (2012) Chromothripsis and cancer: Causes and consequences of chromosome shattering. *Nat Rev Cancer* 12(10):663–670.
8. Nik-Zainal S, et al.; Breast Cancer Working Group of the International Cancer Genome Consortium (2012) Mutational processes molding the genomes of 21 breast cancers. *Cell* 149(5):979–993.
9. Alexandrov LB, et al.; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MMML-Seq Consortium; ICGC PedBrain (2013) Signatures of mutational processes in human cancer. *Nature* 500(7463):415–421.
10. Fischer A, Illingworth CJR, Campbell PJ, Mustonen V (2013) EMu: Probabilistic inference of mutational processes and their localization in the cancer genome. *Genome Biol* 14(4):R39.
11. Bozic I, Allen B, Nowak MA (2012) Dynamics of targeted cancer therapy. *Trends Mol Med* 18(6):311–316.
12. Bock C, Lengauer T (2012) Managing drug resistance in cancer: Lessons from HIV therapy. *Nat Rev Cancer* 12(7):494–501.
13. Bozic I, et al. (2013) Evolutionary dynamics of cancer in response to targeted combination treatment. *eLife* 2:e00747.
14. Read AF, Day T, Huijben S (2011) The evolution of drug resistance and the curious orthodoxy of aggressive chemotherapy. *Proc Natl Acad Sci USA* 108(Suppl 2): 10871–10877.
15. Foo J, Michor F (2009) Evolution of resistance to targeted anti-cancer therapies during continuous and pulsed administration strategies. *PLOS Comput Biol* 5(11): e1000557.
16. Gatenby RA, Silva AS, Gillies RJ, Frieden BR (2009) Adaptive therapy. *Cancer Res* 69(11):4894–4903.
17. Rivoire O, Leibler S (2011) The value of information for populations in varying environments. *J Stat Phys* 142(6):1124–1166.
18. Rivoire O, Leibler S (2014) A model for the generation and transmission of variations in evolution. *Proc Natl Acad Sci USA* 111(19):E1940–E1949.
19. Rouzine IM, Rodrigo A, Coffin JM (2001) Transition between stochastic evolution and deterministic evolution in the presence of selection: General theory and application to virology. *Microbiol Mol Biol Rev* 65(1):151–185.
20. Ewens WJ (2004) *Mathematical Population Genetics* (Springer, New York).
21. Kappen HJ (2005) Path integrals and symmetry breaking for optimal control theory. *J Stat Mech* 2005:P11011.
22. Bertsekas DP (1995) *Dynamic Programming and Optimal Control* (Athena Scientific, Belmont, MA), Vol 1.
23. Bellman RE, Kalaba RE (1964) *Selected Papers on Mathematical Trends in Control Theory* (Dover Publications, New York).
24. Gardiner CW (2009) *Stochastic Methods, Springer Series in Synergetics* (Springer, Berlin), 4th Ed, Vol 13.
25. Cassandra AR, Kaelbling LP, Littman ML (1994) Acting optimally in partially observable stochastic domains. *Proceedings of the Twelfth National Conference on Artificial Intelligence* 94:1023–1028.
26. Swan GW (1990) Role of optimal control theory in cancer chemotherapy. *Math Biosci* 101(2):237–284.
27. De Pillis LG, Radunskaya A (2001) A mathematical tumor model with immune resistance and drug therapy: An optimal control approach. *Comput Math Methods Med* 3:79–100.
28. Ledzewicz U, Schättler HM (2002) Optimal bang-bang controls for a two-compartment model in cancer chemotherapy. *J Optim Theory and Appl* 114(3):609–637.
29. Das Thakur M, et al. (2013) Modelling vemurafenib resistance in melanoma reveals a strategy to forestall drug resistance. *Nature* 494(7436):251–255.
30. Van Kampen NG (2007) *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam), 3rd Ed.
31. Bäuerle N, Rieder U (2013) More risk-sensitive Markov decision processes. *Math Oper Res* 39:105–120.
32. Holohan C, Van Schaeybroeck S, Longley DB, Johnston PG (2013) Cancer drug resistance: An evolving paradigm. *Nat Rev Cancer* 13(10):714–726.
33. Schuh A, et al. (2012) Monitoring chronic lymphocytic leukemia progression by whole genome sequencing reveals heterogeneous clonal evolution patterns. *Blood* 120(20): 4191–4196.
34. Murtaza M, et al. (2013) Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature* 497(7447):108–112.
35. Oesper L, Mahmoody A, Raphael BJ (2013) THetA: Inferring intra-tumor heterogeneity from high-throughput DNA sequencing data. *Genome Biol* 14(7):R80.
36. Roth A, et al. (2014) PyClone: Statistical inference of clonal population structure in cancer. *Nat Methods* 11(4):396–398.
37. Fischer A, Vázquez-García I, Illingworth CJR, Mustonen V (2014) High-definition reconstruction of clonal composition in cancer. *Cell Reports* 7(5):1740–1752.
38. Beckman RA, Schemmann GS, Yeang CH (2012) Impact of genetic dynamics and single-cell heterogeneity on development of nonstandard personalized medicine strategies for cancer. *Proc Natl Acad Sci USA* 109(36):14586–14591.
39. Haeno H, et al. (2012) Computational modeling of pancreatic cancer reveals kinetics of metastasis suggesting optimum treatment strategies. *Cell* 148(1-2):362–375.
40. Yuan Y, et al. (2012) Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Sci Transl Med* 4(157):157ra143.
41. zur Wiesch PA, Kouyos R, Engelstädter J, Regoes RR, Bonhoeffer S (2011) Population biological principles of drug-resistance evolution in infectious diseases. *Lancet Infect Dis* 11(3):236–247.
42. Goldberg DE, Siliciano RF, Jacobs WR, Jr (2012) Outwitting evolution: Fighting drug-resistant TB, malaria, and HIV. *Cell* 148(6):1271–1283.
43. Luksza M, Lässig M (2014) A predictive fitness model for influenza. *Nature* 507(7490): 57–61.
44. Garnett MJ, et al. (2012) Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* 483(7391):570–575.