Research review paper

Vaccine protocols optimization: *In silico* experiences

Francesco Pappalardo\textsuperscript{a,b,⁎}, Marzio Pennisi\textsuperscript{b}, Filippo Castiglione\textsuperscript{a}, Santo Motta\textsuperscript{b}

\textsuperscript{a} Institute for Computing Applications ‘M. Picone’, National Research Council (CNR), Rome, Italy
\textsuperscript{b} University of Catania, Catania, Italy

**Abstract**

Vaccines represent a special class of drugs, capable of stimulating immune system responses against pathogens and tumors. Vaccine development is a lengthy process that includes expensive laboratory experiments in order to assess safety and effectiveness. As the efficacy of a vaccine was demonstrated by biological/chemical investigations and pre-clinical studies, then a major problem is represented by the search for an optimal vaccination dosage. Optimality here assumes the meaning of assuring a high degree of efficacy and safety (lack of toxic or side effects). In lack of quantitative methods, this is usually achieved by a consensus technique, a public statement on a particular aspect of medical knowledge available at the time it was written, and that is generally agreed upon as the evidence-based, state-of-the-art (or state-of-science) knowledge by a representative group of experts in that area. In this article, we focus on the difficult problem of the search for an optimal vaccination dosage in the field of tumor immunology, that is a major issue in biomedical research. This, indeed, represents a first step toward a personalized medicine approach.

© 2009 Elsevier Inc. All rights reserved.

**Contents**

1. Introduction ...............................................................8 2
2. Vaccines in tumor immunology ......................................................8 4
3. Optimal control theory applied to cancer immunotherapy .........................................8 4
4. Genetic algorithms and simulated annealing applied to cancer immunoprevention . . ............................8 6
   4.1. Genetic algorithms approach ....................................................8 8
       4.1.1. The optimization problem .................................................8 8
       4.1.2. Wet biology constrained genetic algorithm search .....................................8 9
   4.2. Simulated annealing approach ...................................................9 0
5. Conclusions ...............................................................9 1
Acknowledgement ..............................................................9 2
References ..................................................................9 2

**1. Introduction**

Drug discovery is the process by which new drugs are designed and tested. The first step in the drug discovery process begins in the laboratory where chemists, pharmacologists and biologists collaborate to identify key factors that play a role in specific diseases. Their major effort is devoted to search for chemical and biological compounds that have drug-like effects. We can identify four steps in the development of new medicines. The sequence starts with the target identification that involves the determination of patterns which are believed to be associated with the disease. The next step is the target validation. In this phase researchers analyze and compare each drug target to confirm and study the effect and the interactions with the diseased organism. This allows to discard targets that do not have the desired effects. The next point is represented by lead identification. Here one or more substances that are believed to have the potential to treat disease are identified. Testing is then performed on each of these molecules to confirm the effect on the target. The lead optimization stage is characterized by the study and test of lead substances in living organisms (*in vivo*) and in cells in the test tube (*in vitro*). The aim is to
compare the pre-selected compounds and to better understand how they are metabolized and ultimately how they affect the whole organism (i.e., its pharmacodynamics).

The process described above is very stringent: only 0.1% of the compounds identified during the discovery process is considered safe for clinical phase I trials in human subjects. Only the 0.02% of the initial pool of compounds is finally approved as a marketed drug for treatment. An average of about 6–8 years pass from the time a drug enters the clinical trials phase until it receives the final approval from the regulatory agencies for sales to the public.

For standard drugs, the dosage is determined by means of the pharmacodynamics (PD) and pharmacokinetics (PK) approach. Dosage is decided for a class of individuals based on PK–PD models and clinical trials. Personalized drug dosage studies have been analyzed for drug cancer administration. The “virtual patient” (Agur, 2006) is an attempt to optimize drug dosage using mathematical models. The output of the virtual patient is a computed optimal dosage which is released to the clinicians for their final decision.

Vaccines represent a special class of drugs. They do not affect the pathogen directly. Instead they stimulate the host to generate a pathogen-specific immune response. Vaccine development is a lengthy process which includes expensive laboratory experiments to determine their effectiveness and safety.

Nowadays computational models have been recognized as relevant for the understanding of biological systems. In particular, models are suitable for guiding biology from a qualitative to a quantitative, thus predictive, science. Pharmaceutical companies are starting to use models to optimize/predict therapeutic effects at the organism level, suggesting that computational biology can effectively play a key role in this field (Kumar et al., 2006).

One of the most exciting challenges is represented by the simulation of the immune system. The immune system represents one of the most complex biological systems. It is, in fact, an adaptive learning system which operates at multiple levels (molecules, cells, organs, organisms, and groups of organisms). Immunological research, both basic and applied, needs to deal with this complexity.

Computational immunologists increasingly use mathematical modeling and computer simulation to study the immune system and the immune responses to different pathogens. Thus, quantitative models that appropriately capture the complexity (both in the architecture and the function) of the immune system are an integral component of the personalized medicine efforts. In silico models of the immune system can provide answers to a variety of questions, including understanding the general behavior of the immune system, the course of disease, the effects of treatments, the analysis of cellular and molecular interactions, and eventually the simulation of laboratory experiments.

From the methodology point of view, we categorize the models into ab initio (theoretical) models that are based on established laws and knowledge of the domain known as first principles, and statistical models that are derived from a limited number of observations or data points where a detailed description of the system is lacking (Pappalardo et al., 2008b).

Theoretical models can be further classified in continuous models or discrete models according to the underlying mathematical framework. Continuous models use the framework of both immunological theories (Perelson and Weisbuch, 1997) (idiotypic network and clonal selection theories), while discrete models mostly use the theory of the clonal selection. From the computational point of view, we can classify the simulation methodologies into mathematical modeling by differential calculus and by discrete methods like Cellular Automata (CA) or Agent-Based models, which is becoming the main tool to perform complex simulations of the immune system (Cohen, 2007; Louzon, 2007).

Differential equation models capture the general behavior of the system and the adjustment of the global parameters of the model can be “relatively” easy. However, given the limits of the analytical treatment these models are limited to a specific observable phenomena, and therefore they often miss to capture the complex interplay of various factors influencing the observed behavior. Moreover differential equation models deal with averaged continuous quantities whereas the immune system is composed by discrete uniquely identified cells. A more detailed review of earlier applications of differential equations for modeling the immune system can be found in Perelson and Weisbuch (1997).

Cellular Automata and Agent-based models are fully discrete dynamical systems that are better suited for computer simulations of biological systems (Stauffer and Pandey, 1992; Cohen, 2007). These modeling paradigms can be precisely tuned to mimic the behavior of the real system (Manneville et al., 1989).

The initial idea of using a discrete automaton in theoretical immunology is due to Kaufman et al. (1985). In that model, various cellular populations are represented by Boolean values and interactions among them are coded by simple rules.

While CA as a modeling paradigm has a relatively long tradition, the agent-based methodology is more recent due to the increasing power of computers enabling large and complex agents to be represented in memory. As in CAs, there are rules governing the interactions but in ABMs these rules are sophisticated algorithms rather than simple Boolean functions. Many of the recent computational models of the immune system are agent-based systems and a large number of the current implementations have been directly or indirectly inspired by the early immunological automaton proposed in Celada and Seiden (1992) whose goal was to build a general immune system simulator.

Both mathematical and computational models have, as everything in this world, both advantages and pitfalls. A computational model uses a formal description in which the primary semantics is operational; this means that a precise and finite sequence of instructions can be executed by an abstract machine, like a computer. On the other hand, a mathematical model differs in the description that uses a denotational semantics. The denotation is made by equations that describe a relationship between quantities and how they change over time. There is an entire sub-field of computer science that studies the relationships and differences between computational (operational) and mathematical (denotational) views of a system. Computational models capture emergent phenomena, provide a natural environment to study the system and new biological knowledge can be easily added (flexibility). However, they don't allow an easy adjustment of parameters and it is not possible to use “exact” methods to solve the model. These two aspects are instead benefits of mathematical models, that are indeed able to capture the general behavior of the modeled system. The pitfall of these kinds of models is that they capture only specific observable functions and are unable to capture the complex interplay of various factors. Despite these considerations, it is worth to note that a third modeling methodology can be used and may be useful to cover the pitfalls of the above described modeling techniques: hybrid models, based both on mathematical and computational models, can be developed and can be used in a synergic way in order to obtain good results.

In immunoinformatics, a common optimization problem is the search for the optimal vaccination schedule of a certain therapeutic agent in the attempt to cure or at least reduce the pathological burden in virtual patients.

In real life, when a newly designed vaccine is ready to be administered for the first time in vivo, whether to mice or to humans, the injection schedule is designed empirically, using a combination of immunological knowledge, experience from previous vaccine endeavors, and practical constraints. In subsequent trials, the schedule of vaccinations is then refined on the basis of the protection elicited in the first batch of subjects and their immunological responses, e.g., kinetics of antibody titers, cell mediated response and so on.

The problem of defining the optimal schedules is of utter importance in cancer immunopreventive approaches, which requires a sequence of vaccine administrations to keep a high level of protective immunity...
against the ever growing malignancy over very long periods, ideally for the entire lifetime of the host.

When compared with this kind of optimization problems, one can resort to different techniques, according to the underlying computational model: stochastic optimization algorithms can be applied when discrete models such as ABMs are used; optimal control theory can be employed when continuous models, i.e., differential equation models, are utilized.

In what follows we focus on the difficult problem of the search for an optimal vaccination dosage in the field of tumor immunology that represents a major issue for biomedical research in particular for its social and economic impact. Then, we illustrate the optimization techniques we have developed to cope with this problem.

The plan of the paper is the following. Section 2 introduces the biological scenario and the state of the art of vaccines in tumor immunology. Section 3 shows how the use of the mathematical theory of optimal control can be used for optimizing immunotherapy vaccine schedule/dosage based on dendritic cells. Section 4 devotes to the use of artificial intelligence methods for finding optimal schedules in cancer immunoprevention vaccines. Finally Section 5 gives conclusions and perspectives.

2. Vaccines in tumor immunology

The efficacy of vaccines have been historically proved: they save more lives than any other drug. However, despite this, there is not a clear understanding of how vaccines work. Nevertheless we are here concerned with another question: once the efficacy of a vaccine has been demonstrated by biological/chemical investigations and pre-clinical studies, how can we determine what is the optimal vaccination dosage, where the optimal means high level of efficacy and safety (lack of toxic effects or side effects) among the population of the treated patients. In lack of quantitative methods, this is usually achieved by a consensus technique, a public statement on a particular aspect of medical knowledge available at the time it is written, and that is generally agreed upon as the evidence-based, state-of-the-art (or state-of-science) knowledge by a representative group of experts in that area.

Recent research (Guinn et al., 2007) demonstrated that the vaccine approach may also be useful in the prevention and treatment of cancer (tumor immunology). It is known that the immune system eliminates most of the cancer cells. Those that are not recognized escape immune surveillance, leading to tumors. Tumor vaccines can thus be used to stimulate an immune response against poorly immunogenic tumor variants. In few words, the ultimate goal of tumor immunology is to understand the interactions between tumor and immune system cells, and to devise immune based approaches to fight cancer.

The use of cytotoxic T cells (CTLs), dendritic cells (DC) and antibodies, actually represent well-known approaches in cancer immunotherapy (Begley and Ribas, 2008). The use of anti-idiotypic (Id) antibodies as vaccines to stimulate immune system response against tumors have been demonstrated effective in preventing tumor growth and curing mice with established tumors (Bhattacharya-Chatterjee et al., 2002). Several monoclonal anti-Id antibodies that have the appearance of distinct human tumor-associated antigens (TAA) have been developed and tried in the clinic, demonstrating good results. Indeed the efficacy of these vaccines will depend on the results of several Phase III clinical trials. Numerous studies in mouse tumor models have shown that DCs pulsed with tumor antigens can induce protective and therapeutic anti-tumor immunity (Nestle, 2005). It is, however, worth to mention that the complexity of the DC system requires rational manipulation of DCs to achieve protective or therapeutic immunity. Further research is needed to analyze the immune responses induced in patients by distinct ex vivo generated DC subsets that are activated through different pathways. These ex vivo strategies should help to identify the parameters for in vivo targeting of DCs. One major stumbling block for cancer immunother-

apy is that it is usually tested in advanced cancer patients. This is a reasonable setup for early trials of new cytotoxic drugs that directly hit cancer cells. In the experimental environment the trend to move immune intervention to the early stages of tumor development is taking even more extreme incarnations (cancer immunoprevention). Recently it has been shown that prophylactic vaccines administered to transgenic mice prone to cancer development can completely prevent tumor onset and restore a normal life expectancy (Lollini et al., 2006a). Even though prophylactic cancer vaccines are still far from human application, this opens up an entirely new perspective in cancer prevention, leading to a future in which vaccines will equally contribute to the prevention of infectious diseases and cancer.

3. Optimal control theory applied to cancer immunotherapy

The theory of optimal control has been already applied to cancer. For example Acharya and Sundaresan (1984) discuss about optimal drug delivery, Swan (1990) gives a review of the ways in which optimal control theory is applied to growth kinetic models, cell cycle models, and a classification of “other models” together with suggestions for designing better chemotherapy strategies. de Pillis and Radumskaya (2001) also face the problem of administration of chemotherapy but considering the interesting question of immune resistance.

As already stated above, in the specific case of cancer immunotherapies the therapeutic agent is something able to elicit an immune response against the malignancy. The typical question then is how to decide how much of the therapeutic agent to inject and when, so to have a protective and prolonged immune response. This question can be addressed in mathematical terms in two stages: the first one constructs a mathematical model describing the cancer-immune interaction and secondly one applies the theory of optimal control to determine when and to which extent to stimulate the immune system.

In Castiglione and Piccoli (2007), a method that can be applied to find the optimal protocol in a variety of clinical problems where the kinetics of the drug or treatment and its influence on the physiologic/pathologic functions have been described by a system of ordinary differential equations is shown.

Here we deal with dendritic cell transfection immunotherapy. Dendritic cell transfection is the practice of cultivating autologous dendritic cells (i.e., previously extracted from the same patient), together with some known tumor-associated-antigen (TAA) and then inject them back into the patient. The resulting vaccine made by autologous TAA-loaded dendritic cells is called dendritic cell vaccine (DCV). The idea is that the immune system, confronted with such amount of tumor-antigen, starts to mount a response against it, in place of an otherwise weak or completely absent response. In fact, as a side effect of the immune response against the DCV, the immune system will eventually recognize the same TAA molecule on tumor cells and kills them. In Castiglione and Piccoli (2006) the main concern was to develop a generic procedure to optimize the time/dosage dependent administration of the TAA so to reduce or, in the best case, to eliminate the tumor mass.

Similar works have been previously published (Martin, 1992; Kirschner and Panetta, 1998; Swiemial et al., 2003; Burden et al., 2004; Fister and Donnelly, 2005). In many of these references, the treatment is thought as a continuous process, thus the used techniques are those typical of optimal control problems in continuous time. On the contrary in Castiglione and Piccoli (2006) discrete injection times were introduced.

Here we describe the work in Castiglione and Piccoli (2007) where, besides the vaccination schedule, also the dosage as control variables was considered. The ODE model of the tumor-immune interaction is quite simple (Castiglione and Piccoli, 2006) and is based on the assumption that dendritic cells, as the source of TAA, are introduced externally to ignite the immune response against themselves and, as a side effect, also against the tumor cells. In fact, the clone expansion of cytotoxic T cells able to recognize the TAA-loaded by DCs, also favor
tumor killing since cancer cells naturally display the same TAA on their cell surface. In mathematical terms, this means:

\[ H = a_0 + b_0 DH (1 - H - f_0) - c_0 H \]  

(1)

\[ \dot{C} = a_1 + b_1 I (M + D) (1 - C - f_1) - c_1 C \]  

(2)

\[ \dot{M} = b_2 M (1 - M - f_2) - d_2 MC \]  

(3)

\[ \dot{D} = -d_3 DC + u \]  

(4)

\[ I = b_4 DH - e_1 I C - c_4 I \]  

(5)

where \( H \) are the tumor-specific CD4 T helper cells, \( C \) are the tumor-specific CD8 cytotoxic T cells, \( M \) are the cancer cells that expose the TAA, \( D \) are the mature dendritic cells loaded with the TAA and \( u \) is the control, i.e., the injection rate of dendritic cells. \( I \) is the IL-2 secreted by \( H \) and responsible for \( T \) cell growth.

In this model the time resolution is of \( 1 \) h. Moreover, one considers only the dynamics of those clones of cells that actually recognize the TAA, neglecting the effect of cross-reactivity of other clones. Finally, the model is meant to be valid in the range of the tumor mass for which the effects of immune escape, immune down-regulation or vascularization are still negligible (Preziosi, 2003). Starting from the set of values used in Kirschner and Panetta (1998) and by tuning the system to reproduce qualitatively the dynamics of the tumor-immune competition the parameters reported in Castiglione and Piccoli (2007) have been sorted out.

The treatment is modeled as a control problem \( x = F (x, u) \), where \( x = (H, C, M, D, I) \in \mathbb{R}^5 \) is the state variable measuring the cell populations, \( u \in U \) is the control, i.e., measures the treatment effects. Since the ultimate goal is both to reduce the tumor mass at the end of treatment period and to keep it under control during the treatment period (i.e., below a certain threshold), one considers an optimal control problem with a cost including a term measuring the tumor killing since cancer cells naturally display the same TAA on their cell surface. The state variable measuring the cell population as a function of time. Thus the total quantity injected is \( \int a_0(s) ds \).

Having defined \( Q_i \) as the total injected quantity of dendritic cells at the \( i \)-th vaccination, to impose the constraint on the periods of control activity to be short and well separated, one simply adds the term \( \sum Q_i \) to the cost where the \( Q_i \)'s are now new control variables. Analogously to the case of continuous control, the analytical solution of the hybrid control problem requires the use of the Maximum Principle.

Then a control procedure that consists in \( N \) vaccinations inoculated according to a schedule \( S = \{ (t_i, Q_i); i = 0, \ldots, N \} \) is considered. Let \( S \) be the space of schedules, then for every \( S \in \mathcal{S} \) we define \( u_S \) to be the corresponding control \( u_S(t) = \sum_{i}^{N} Q_i (t-t_i) \chi(t, t_i) \) where for simplicity \( \int a_0(s) ds = 1 \). The control \( u_S \) corresponds to \( N \) vaccine administrations that occur at times \( t_i \) with injected quantities \( Q_i \). Finally one set: \( U = \{ u_S : S = S \} \).

The problem described above can be then formulated in mathematical terms. Consider the cost obtained summing up the cost (6), (7), with the cost \( \sum Q_i \) and get the problem:

Problem P

Given the initial condition \( x_0 \) determine a schedule \( S = S \) of \( N \) injections so that the control \( u_S \) and the trajectory \( x_S \) of \( x = f(x(t)) + u_S(t) \) satisfies the minimum of the cost, that is:

\[ \min_{S \in \mathcal{S}} \left\{ \int_{t_0}^{t_f} L(x(t), u(t)) dt + \psi(x(T, u)) \ | \ x(0) = x_0 \right\} \]

(10)

where \( w_i \) are weights that can be chosen to assign more importance to this or that cost and \( M_t \) is the tumor mass.

In Castiglione and Piccoli (2007) it is shown that the ingredients for the optimization algorithm for the numerical solution of problem (P) are the computation of \( \frac{\partial \psi}{\partial t_i} \) and \( \frac{\partial \psi}{\partial Q_i} \) (the details are not reported here for simplicity). The optimization algorithm consists of the following procedure:

**step 0:** initialize variables

- \( T \in \mathbb{R}^+ \), \( M_0 \in \mathbb{R}^+ \), \( N \in \mathbb{N}^+ \), \( \mathcal{Q} \in \mathbb{R}^+ \)
- \( x_0 = (H(0), C(0), M(0), D(0), I(0)) \in (\mathbb{R}^+)^5 \)
- \( S_0 = \{ (t_i, Q_i) \}_{i=1}^{N} \)

FOR \( n = 0 \) TO optSteps DO

**step 1:** Solve the system, i.e., \( x_i = \text{RungeKutta}(x_0) \);

Solve the variational equations as in (Castiglione and Piccoli, 2007);

**step 2:** Compute \( \frac{\partial \psi}{\partial t_i} \) and \( \frac{\partial \psi}{\partial Q_i} \);

**step 3:** Update the schedule by and optimization method, i.e., \( \forall i \)

\[ t_i^{n+1} = t_i^n + h_t \frac{\partial \psi}{\partial t_i} \]

\[ Q_i^{n+1} = Q_i^n + h_Q \frac{\partial \psi}{\partial Q_i} \]

for small parameters \( h_t < 0 \) and \( h_Q < 0 \)

END DO

In other words, fix the time horizon \( T \), the maximum allowed value of the tumor mass during treatment \( M_{max} \), the number of vaccine
administrations $N$, the vaccine quantity $Q$, an initial value $x_0$ of cell populations and an initial schedule $S_0$ (this can be taken at random). Then itterates the following steps: solve the system (1)–(5) with $x_0$ as the initial value via the fourth-order Runge–Kutta integrator generating an approximation of the trajectory $x(t)$ at the same time solve the variational equations and compute the derivatives of the extended cost $\phi$ (in Eq. (10)) with respect to the injection times $t_i$ and the vaccine quantities $Q_i$ (details can be found in Castiglione and Piccoli (2007)); update the schedule by the steepest descent (or other optimization) method, for small parameters $h_1<0$ and $h_2<0$.

Summarizing, the question of how to determine the best time/disposal allocation of the injections of dendritic cell vaccination within the therapeutic period has been transformed in a control problem where the cost function is the sum of the tumor mass at the end of the therapy $M(T)$, the integral of the tumor mass exceeding a certain level $M_{\text{max}}$ and the total quantity of vaccine injected.

The hybrid approach was used and the control set simplified to better capture the characteristic of the vaccine administration procedure and reduce the problem to a finite dimensional one. In particular, we used tools of optimal control to compute the gradient of the cost function with respect to the schedule. The latter is obtained via the solution of a generalized variational equation.

The optimal schedule found by means of the above algorithm is characterized by three general, thus interesting, issues: the first few injections are concentrated at the beginning of the treatment period (some vaccinations are glued together) while the other ones are distributed more or less at equal distances. Moreover, the very first vaccination is the most consistent in terms of dosage, so to readily diminish the tumor mass whereas the remaining ones have the purpose of keeping the tumor under control. Nevertheless it is noteworthy that the result has been found by imposing general and reasonable constraints on the vaccine administration protocol, something that is part of the common clinical practice, as for example, the adoption of treatment interruptions (drug holidays) to lessen the burden of the patient.

4. Genetic algorithms and simulated annealing applied to cancer immunoprevention

In immunoprevention the vaccine action needs to be efficient in the very first appearance of the tumor cells. Those vaccines are not designed, and then not effective, if a solid tumor is already established. The vaccine action is to induce an immune response that controls the number of tumor cells and therefore prevents the development of a solid tumor mass. A successful vaccination schedule should then be able to prevent the formation of the solid tumor for the entire lifetime of the host.

An example of an immunopreventive vaccine is represented by Triplex vaccine (Lollini et al., 2006a), designed to be effective against mammary carcinoma. It is based on multiple immune signals in the same vaccine. The Triplex vaccine combines the target antigen with two “adjuvant” stimuli, interleukin 12 (IL-12) and allogeneic Major Histocompatibility Complex (MHC) molecules. The main purpose of IL-12 is to enhance antigen presentation and Th cell activation in response to the antigen. Allogeneic MHC molecules stimulate multiple T cell clones and cause a broad production of immunostimulatory cytokines that amplify immune responses. The first formulation of the Triplex vaccine consisted of MHC–allogeneic mammary carcinoma cells expressing HER-2/neu and of recombinant IL-12. Subsequently the need for IL-12 administration was bypassed through the transduction of vaccine cells with IL-12 genes.

Repeated vaccination of HER-2/neu transgenic mice with the Triplex vaccine starting at an early age (6 weeks) completely prevented the onset of mammary carcinoma for up to one year of age (Nanni et al., 2001). Complete immunoprevention of mammary carcinogenesis was obtained with a chronic schedule of vaccinations, i.e., a cycle of two vaccinations per weeks for two weeks followed by two weeks of rest, usually repeated for one year.

The major issue still unresolved with the Triplex vaccine is whether or not the Chronic schedule is the minimal set of vaccination yielding complete, long-term protection from mammary carcinoma. Shorter vaccination protocols failed to prevent cancer, but between shorter protocols and the Chronic one there still is an infinite set of schedules that might yield complete protection with significantly fewer vaccinations. From an experimental point of view this would require numerous sets of experiments each lasting one year, a feat that would discourage any wet biology team. It is worth to point out here that the matter at stake is not a problem of biological laziness, but rather of translational research. The goal of this type of research is not just to prevent mouse tumors, but to devise strategies that could be implemented in humans (Lollini et al., 2006a). The Chronic protocol would lead to human protocols entailing frequent vaccinations for the entire lifetime of subjects at risk of cancer, making for a very cumbersome and unpractical attempt at translating promising pre-clinical results. Biological experiments based on the results of the simulations described in Lollini et al. (2006b) could lead to ease clinical translation.

To attack the problem of finding an optimal vaccination schedule for the Triplex vaccine, the first step is to have a computational model able to reproduce pre-clinical data on mice. SimTriplex (Pappalardo et al., 2005) is an ABM that describes all the relevant processes of the “immune system–mammary carcinoma” competition by means of rules derived from biological experiences and knowledge. Such a paradigm has been used to model other immune system related pathologies, like atherosclerosis (Pappalardo et al., 2008a). This model and simulator includes all the most relevant entities (e.g., immune system cells, cancer cells, vaccine cells) and processes (e.g., recognition, duplication, immune memory, humoral and cell mediated cytotoxicity) needed to reproduce the immune response induced by the Triplex vaccine.

The model incorporates, at its current stage the principal core facts of today’s immunological knowledge, e.g. the diversity of specific elements, human leukocyte antigen (HLA) restriction, clonal selection by antigen affinity, thymic education of T cells, antigen processing and presentation (both the cytosolic and endocytic pathways are implemented), cell–cell cooperation, homeostasis of cells created by the bone marrow, hypermutation of antibodies, maturation of the cellular and humoral response and memory. Important immune system functions such as thymic selection, representation of the repertoire, affinity maturation and antibody hypermutation are also implemented. Major classes of cells of the lymphoid lineage (B and T lymphocytes) and some of the myeloid lineage (macrophages and dendritic cells) are represented. For what concerns molecules, the model distinguishes between simple small molecules like interleukins or signaling molecules in general and more complex molecules like immunoglobulins and antigens, for which we need to represent the specificity. We only represent interleukins 2 and 12. Interleukin 2 is necessary for the development of T cell immunologic memory, one of the unique characteristics of the immune system, which depends upon the expansion of the number and function of antigen-selected T cell clones. Interleukin 12 is represented as the Triplex vaccine uses that as an adjuvant. For what is related to the immunoglobulins, we represent only type IgG. This is just because at the actual state we don’t need to represent other classes of Ig and because IgG is the most versatile immunoglobulin since it is capable of carrying out all of the functions of immunoglobulins molecules. Moreover IgG is the major immunoglobulin in serum (75% of serum Ig is IgG) and IgG is the major Ig in extra vascular spaces. Looking at Fig. 1, at the same level of entities we find immune system activities. They include both interactions and functions. Functions refer to the main immune system tasks. In particular SimAthero takes care of the diversity of
Fig. 1. SimTriplex detailed graphical representation showing all the immune system entities (both cellular and molecular), function and activities modeled.
specific elements, major histocompatibility classes restriction, clonal selection by antigen affinity, thymus education of T cells, antigen processing and presentation (both the cytosolic and endocytic pathways are implemented), cell–cell cooperation, homeostasis of cells created by the bone marrow, hypermutation of antibodies, cellular and humoral response and immune memory. SimTriplex represents receptors and ligands as bit strings and use a string matching rule to model affinity. This clever idea was introduced by Farmer et al. (1986) as a way to perform calculations for determining molecular complementarity and predicting the optimal size of an epitope. From immunology, we know that binding is a threshold effect consisting of two components: the affinity of a single receptor and ligand, and the total binding, or avidity of multiple binding pairs. Binding is modeled by a string matching rule by counting the number of positions in the string at which the symbols are complementary (known as Hamming distance). Repertoires are represented in the model as sets of strings. This fundamental modeling abstraction ignores nearly all of the physical and chemical details that determine receptor/ligand interactions. By adopting bit strings, many binding events can be simulated quickly, making it feasible to study large-scale properties of the immune system. Although character strings are a poor representation of the reality, they produced accurate models when benchmarked to experiment, suggesting that the abstraction captures important features of receptor/ligand binding.

In particular, specificity is implemented in SimTriplex by a bit-string polyclonal lattice method. Bit-string refers to the way the molecules and the specificity among molecules is represented, polyclonal indicates that more clones of different specificity of lymphocytes are represented and lattice means that we use a discrete lattice to represent the space, that is, the space is discrete. The set of lymphocyte's receptors is represented by a bit-string of length \( L \) which then forms the so-called shape space. A clonal set of cells is characterized by the same clonotypic receptor, i.e., by the same bit-string of length \( L \). The potential repertoire of receptors scales as \( 2^L \). The receptor–co-receptor binding among the entities are described in terms of matching between binary strings with fixed directional reading frame. Bit-strings represent the generic 'binding site' between cells (through their receptors) and target molecules (through peptides and epitopes).

The simulator takes care of the main interactions that happen during an immune response stimulated by Triplex vaccine against mammary carcinoma.

Fig. 1 shows a detailed representation of SimTriplex connection with cellular and molecular immunology. Some details were omitted due to complexity of the figure.

A validated simulator will reasonably reproduce, in the validation range, the immune response activated by a vaccination protocol, so to reproduce different vaccination schedules in silico and to search for the schedules with a minimal number of vaccine administrations still capable to prevent carcinoma-in-situ formation.

Different strategies to approach this optimization problem have been tried. The first attempt was made by “trial and error”. One set successively repeating cycles of injections at different stages of the virtual mouse age and the simulator was used to determine the survival rate of vaccinated mice. In this way an effective schedule of only 44 vaccinations, that is 27% less than the standard Chronic protocol was found (Motta et al., 2005).

A second search strategy was based on Genetic Algorithms, which are well-known general purpose optimization methods. The third strategy was based on the Simulated Annealing approach. In what follows we describe those last two approaches and the results obtained. Both optimization methods use SimTriplex to validate their protocol suggestions.

### 4.1. Genetic algorithms approach

Genetic algorithms (GAs), conceived by Holland in 1960, are a particular class of evolutionary algorithms (Corne et al., 1999) that use techniques inspired by evolutionary biology. Common applications of GAs are discrete optimization problems of multivariate functions, \( f(x) \), where \( x \) is the vector of variables \( x_1, x_2, \ldots, x_n \). A GA is a method for moving from one set \( \{x\} \) to a new one by using a kind of natural selection rule together with the genetic-inspired operators of crossover, mutation and so on.

In the standard GA language the set \( \{x^k, k = 1, \ldots, m\} \) is referred to as a population and each \( x_i \) is a chromosome. A chromosome \( x_i \) is an \( n \)-dimensional string, usually a binary string; each element of this string is referred to as a gene; each gene can have different instances, i.e., alleles.

We present an application of GA to approach the problem of searching optimal schedules. The same approach has been successfully applied to optimize drug dosage in HIV related HAART therapy (Castiglione et al., 2007). For this purpose we use a GA search on a model based approach (Pappalardo et al., 2006a). In what follows the key points are: i) a genetic algorithm can be used in vaccine protocol design; ii) a genetic algorithm can be “driven” to achieve real biological constraints; and iii) a quantitative approach can be envisaged, as a medium term research goal, for vaccine schedule design.

#### 4.1.1. The optimization problem

In order to translate the biological concept of vaccine effectiveness one considers a time interval \([0, T]\) in which the action of the vaccine on a set of mice \( S \) is studied. This can be either the lifetime of the mouse or the time-length of the in vivo experiment. Then the given time interval is discretized in \( N \) equally spaced subintervals of width \( \Delta t \), i.e., \( \{t_1 = 0, t_2, \ldots, t_N = T\} \). The time interval \( \Delta t \) corresponds to possible vaccine administrations, e.g., every 8 h.

Let \( x = [x_1, x_2, \ldots, x_N] \) be a binary vector representing the sequence of vaccine schedule where \( x_i = 0 \) means respectively administration/no administration of the vaccine. The label \( i \) represents the time \( t_i \) of the vaccine administration and \( t_N \) the end of the vaccination period. The number of vaccine administrations is given by \( n = \sum_{i=1}^{N} x_i \).

Let \( \tau(x, \lambda) \) be the time of carcinoma-in-situ (CIS) formation. This is a function of the vaccination schedule \( x \) administered to the mouse \( j \) of size \( S \) and a parameter \( \lambda_j \) which describes the biological diversity. The vaccine will be effective if \( \tau > T \).

Then the unconstrained optimization problem is to find \( \mathbf{x} \) such that:

\[
\tau(x, \lambda_j) = \max\{\tau(x, \lambda_j)\}
\]

\[
(x) = \min\{n(x)\}
\]

This is a multi-objective discrete optimization problem. The search space \( D \) has cardinality \( 2^n \). For \( T = 360 \) days, and \( \Delta t = 24 \) h the cardinality is \( 2^{360}-10^{100} \) which means there is no hope for an exhaustive search.

The function \( \tau \) depends on the action of the stimulated immune system. Obviously such a function cannot be expressed in an analytical form and \( \tau(x, \lambda) \) can only be computed through an ad hoc simulator. In our case this is the above mentioned SimTriplex simulator.

Given a multi-objective optimization problem, max \( g_1 \) and min \( g_2 \), if there is no conflict between the two objectives, the standard technique to find a solution is to combine them into a single-objective problem for a new function \( f(g_1, g_2) \). The simplest case is to consider a linear combination \( f = \alpha g_1 + \beta g_2 \), where \( \alpha \) is a weight parameter.

A linear combination does not fit the problem. Then, one chooses:

\[
f_1(n(x), \tau(x, \lambda)) = \frac{n(x)}{\tau(x, \lambda)}
\]

and problem (11) and (12) reduce to: find \( \mathbf{x} \) such that:

\[
f_1(n(x), \tau(x, \lambda)) = \min\{f_1(n(x), \tau(x, \lambda))\}
\]
optimization problem. Let \( f(n, \tau, \ldots) \) which just takes into consideration Eq. (11) will give high rank to schedules with many vaccine administrations. On the other hand, if one considers just Eq. (12), then schedules with a lower number of vaccine administrations will have higher rank. To take into consideration both Eqs. (11) and (12) the fitness function must be, at least, a two-variables function of the type \( f(n, \tau, \ldots) \) which satisfies the following two properties:

\[
f(n, \tau, \ldots) < f(n, \tau', \ldots) \quad \text{if } \tau > \tau' \\
f(n, \tau, \ldots) > f(n', \tau, \ldots) \quad \text{if } n > n'
\]

(15) (16)

The function (13) satisfies properties (15) and (16) so, as expected, \( f(n, \tau, x, \lambda_i) \) is also the fitness function.

This problem has been solved in Pappalardo et al. (2006a; Lollini et al., 2006b). However the solution was biologically unsatisfactory because:

i) the optimal schedule \( X \) yields very high peaks in cancer cells. Even if a CIS is not yet formed, a high number of cancer cells may induce, by overstimulation, an anergic state of T lymphocytes, depleting in this way the immune system response and enhancing the risk of carcinogenesis.

ii) due to biological variability the optimal schedule found for mouse \( j \) is able to prevent CIS formation only in a small fraction (~20%) of mice of the set \( S \). Therefore the schedule found for a single mouse is not immunologically effective since protection does not generalize to more than one mice.

Following the medical experts advice, to overcome i) one needs to add to the problem (14) two constraints on the maximum number of allowed cancer cells in order to prevent T lymphocytes to go into the anergic state. Let \( M_1(x) \) and \( M_2(x) \) respectively be the maximum number of cancer cells in \([0, T_{in}]\) (cellular-mediated controlled phase) and in \([T_{in}, T]\) (humoral-mediated controlled phase) where \( T_{in} \approx T/3 \). Let \( \gamma_1 \) and \( \gamma_2 \) be, respectively, cancer cells threshold in \([0, T_{in}]\) and in \([T_{in}, T]\). Our optimization problem (14) now becomes:

\[
\begin{align*}
\min & \quad f_i(n(x), \tau(x, \lambda_i)) \\
\text{subject to :} & \\
M_1(x) & \leq \gamma_1, \quad t \in [0, T_{in}] \\
M_2(x) & \leq \gamma_2, \quad t \in [T_{in}, T]
\end{align*}
\]

(17)

Notice that, at variance with usual methods in continuous optimization, the optimal search performed using GA translates the constraints into penalties of the fitness function. Note that solutions which do not satisfy the constrains are still accepted but with lower ranking and therefore the optimal solution found may still not satisfy all the constrains.

To solve the second problem, ii), one needs to reformulate the optimization problem. Let \( \{j_1, j_2, \ldots, j_m\} \subset S \), with \( m = 8 \), a random chosen subset of in silico mice. The optimization problem is to find \( X \) such that:

\[
\begin{align*}
\min & \quad f_i(n(x), \tau(x, \lambda_{j_1})) \\
\min & \quad f_i(n(x), \tau(x, \lambda_{j_2})) \\
\cdot & \\
\min & \quad f_i(n(x), \tau(x, \lambda_{j_8})) \\
\text{subject to :} & \\
M_1(x) & \leq \gamma_1, \quad t \in [0, T_{in}] \\
M_2(x) & \leq \gamma_2, \quad t \in [T_{in}, T]
\end{align*}
\]

(18)

This is again a multi-objective optimization problem that one can reduce to a standard problem defining the new objective function as a linear combination of \( f_{j_1}, \ldots, f_{j_n} \), that is:

\[
f(n, \tau_{j_1}, \tau_{j_2}, \ldots, \tau_{j_8}) = \sum_{k=1}^{8} \alpha_k f_{j_k}(n, \tau_{j_k})
\]

(19)

where all \( \alpha_k = 1 \) as the mice are equiweighted.

This is again a multi-objective optimization problem. This is reduced to a standard problem like in (18) finding the minimum of (19) subject to the above described constrants. The results are discussed in Lollini et al. (2006b).

As the computed optimal vaccination schedule must be verified in vivo, additional constrains required by wet-lab procedures must be considered.

The most important of these requirements is the regular weekly administration (only Monday and Thursday). This is considered a very intensive vaccination schedule from an immunological point of view and it is also labor-intensive, because it mandates the preparation of fresh cells that make the vaccine twice a week.

4.1.2. Wet biology constrained genetic algorithm search

In the genetic algorithm, a chromosome in the chromosomes' population represents a vaccine schedule. Fig. 2 details the representation used by the GA and its translation in a ready wet-lab schedule.

In vivo experiments last usually for 400 days. Taking into account that the time step of the simulator is 8 h of real life this leads to a mouse with 1200 alleles. Without any constraint on vaccine administration each allele can take the two instances 0/1 (Pappalardo et al., 2006b).

The wet-lab requirements drastically reduce the number of alleles which can take two instances. These constraints cannot be translated into penalties in the fitness function because they must be completely satisfied. Therefore a different strategy to constraint the optimal search is needed.

As pointed out these constraints drastically reduce the search space as a large number of alleles are forced to have the value 0. The number of such alleles can be easily computed as 1200 time steps corresponds to 400 days: 400 days are 57 weeks plus 1 day. As the mice enter into the experiment on Monday, this leads to 115 days in which the vaccine can be administrated. Thus the cardinality of the new search space \( D \) reduces \( 2^{115} \) (while the complete search space \( D \) was \( 2^{1200} \)).

Since the simulator which computes the function \( \tau \) runs with a time step of 8 h we need to map the new search space into the old one. A function which maps 115 days schedule into 1200 time steps schedule was created and applied in the modified GA. The function is shown below:

\[
M(i) = 9i + 3 \left( \binom{i}{2} + 1 \right)
\]

Fig. 2. Correspondence between vaccinations and the elements of a genetic algorithm. A complete vaccination schedule was encoded into a bitstring. Each bit represents an 8 h time step in which a single vaccination can be administered (bit = 1) or not (bit = 0). In genetic algorithm parlance each bit is a gene and the entire bitstring a chromosome.
where $i \in [1,115]$ represent the available days (genes) for the vaccination administration in the GA. $M(i) \in [1,1200]$ returns the corresponding of the 1200 time steps.

With this function, the genetic algorithm will only search two days within a week, which are suitable for vaccinations under the administration. Using this constraints the in silico experiment found a 37 vaccination schedule with a survival of 88%.

Remarkably, these results are almost equivalent to those produced with the day-unconstrained GA model both in terms of the number of injections required and the survival rate reached.

The successful coupling of the genetic algorithm and the simulator was used as the basis for the in silico design of vaccination protocols to be tested in vivo. Prof. Lollini lab agreed to test in vivo a vaccination schedule produced by the genetic algorithm entailing a 50% reduction in the number of vaccinations in comparison to the Chronic protocol (Fig. 3). This schedule is indicated as “Gen 1.1” in Fig. 3. “Gen 1.1” schedule was obtained removing the last five injections from the GA final suggested protocol. This because biologists believed that these last five injections will not affect survival outcome. The predicted in silico survival percentage was however of 81%. A key issue at this point was the design of appropriate control schedules. We included in the in vivo experiments three controls of known protective potency, namely:

- untreated mice, to define the baseline level of mammary carcinogenesis;
- mice vaccinated according to the Chronic protocol, which represents the current “gold standard” of high protective efficacy in BALB-NeuT mice;
- mice vaccinated according to the “Early” vaccination protocol, which should yield an intermediate survival curve significantly different from those of the two groups above.

A prominent feature of the test schedule produced by the genetic algorithm was the irregular spacing of the intervals between vaccinations over time, with a higher density at earlier time points. Hence this schedule differs from the “Chronic” one in two aspects, i.e. number of vaccinations and intervals. To discriminate between these two variables we introduced a further control group, using the same number of vaccinations as the test schedule, regularly spaced over the same period. In practice this is the protocol a human immunologist would design on a heuristic base with the sole constraint of the total number of vaccinations (designated “Heur 1.0” in Fig. 3). In vivo testing is ongoing and results will be published in due course.

### 4.2. Simulated annealing approach

Another famous stochastic optimization technique particularly suitable for global optimum problems is represented by the Simulated Annealing (SA) (Kirkpatrick et al., 1983). It rose up from observing that there is a deep analogy between Statistical Mechanics and Combinatorial Optimization. There is a direct similarity, indeed, between the “fast warming–slow cooling” process used to give significant strength and hardness properties to solids (annealing) and the framework required for optimization of complex systems.

The Metropolis Algorithm (Metropolis et al., 1953) lies behind the SA. Given a combinatorial optimization problem and an initial configuration of a system with its own energy (i.e., a proper measure of “quality” of the configuration, usually defined as a cost function), the Metropolis Algorithm randomly perturbs this configuration. If the difference in energy, $\Delta E$, between the old and the perturbed configuration is negative, the new configuration indicates a state with lower energy and therefore this state is assumed as the new one. Otherwise the probability of acceptance of the new configuration is given by the Boltzmann factor (a weight factor that determines the probability of a configuration in a multi-state system in thermodynamic equilibrium at given temperature $T$). After a “suitable” number of perturbations, the given configuration should approach thermal equilibrium at temperature $T$ (i.e., a state with minimum energy for the given temperature). The Metropolis algorithm ensures the achievement of the thermal equilibrium if the number of iterations approaches infinity. Since this condition is not satisﬁable, it is common practice to set a maximum number of iterations.

We can regard the SA algorithm (at least in its homogenous form) as a sequence of Metropolis algorithms evaluated at a sequence of decreasing values of the temperature $T$, used as control parameter. If $T$ is lowered sufﬁciently slow, it is possible to demonstrate that the final frozen conﬁguration will have minimal energy and will represent the optimal solution for the problem at hand.

The first step in implementing the SA algorithm for the problem of finding optimal vaccine schedules has been represented by an attempt to find some analogies between the relevant entities of SA (temperature, energy, conﬁguration) and the relevant concepts defining a vaccine protocols (number of injections, distribution of injections, mean given survival).

As previously said, any protocol can be deﬁned as optimal if it is able to guarantee the best population survival rate with the minimum number of vaccine administrations. Given that the experiment lasts approximately 57 weeks and no more than two administrations per week are feasible, any candidate protocol is described in a similar way as in GA, that is, like a binary vector $x$ of cardinality $V = 115$, where the bit position represents the administration time $t$ and the bit value 1/0 represents vaccine, or no-vaccine administration at time $t$. This vector, in conjunction with its number of administrations $n$, represents a candidate configuration.

### Tumor-free mice • Vaccinations

<table>
<thead>
<tr>
<th></th>
<th>Tumor-free mice</th>
<th>Vaccinations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>In vivo</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early</td>
<td>0%</td>
<td>12</td>
</tr>
<tr>
<td>Chronic</td>
<td>85%-100%</td>
<td>60 (100%)</td>
</tr>
<tr>
<td>Heur. 1.0</td>
<td>65%</td>
<td>30 (-50%)</td>
</tr>
<tr>
<td>Gen. 1.1</td>
<td>81%</td>
<td>32 (-47%)</td>
</tr>
</tbody>
</table>

**Fig. 3.** Schedules tested in vivo in BALB-NeuT transgenic mice. Each tick represents one vaccination. The percentages of tumor-free mice were evaluated either in vivo or in silico, as indicated.
The control parameter of the SA algorithm (the temperature $T$) is slowly but constantly lowered in order to reach the minimum energy. Starting from a defined initial number of injections $n_{in}$, the goal is to minimize the number $n$ of injections of a protocol, trying obviously to keep high survival percentages. The most reasonable choice is therefore to couple $T$ with $n$. Gain survival with no injections is obviously impossible. In this case it has been preferable to set a minimal number of injections $T_{min}$.

Note that if the number of vaccine administrations is decreased from $T$ to $T-1$ and the Metropolis algorithm is not able to find a semi-equilibrium configuration, i.e., an acceptable value of survivals at temperature $T-1$ in $j$ iterations, the accepted protocol will be the last found at temperature $T$.

The configuration energy in SA is a slowly stochastically decreasing quantity. A reasonable measure of the energy should be linked to an opportune mean value of the virtual mice sample survival. The concept of energy is therefore defined as a quantity proportional to the harmonic mean $H$ of the survival times $\tau_i$, $H(\mathbf{x}, \lambda^*) = \frac{1}{\lambda_1 + \ldots + \lambda_m}$. As a matter of fact $H$ decreases when the cumulative survival time of the sample increases, in perfect accord with the energy definition. It’s worth to point out that the survival time of the virtual mice cannot be computed directly but is calculated through SimTriplex simulator.

The Metropolis random perturbation of a configuration is defined as a random bits reallocation. It can be further improved using biological knowledge. As mice survival is a major objective, in moving from $x_i$ to $x_i’$, it is possible to re-adjust random bits reallocation transferring some “1” at a suitable time $t<\min(\tau_i), i=1,k$.

Finally, the SA algorithm for the protocol optimization problem can be defined as follows:

**SA for the Optimal Protocol Search Problem**

$T = \text{a defined initial number of injections } n_{in}$
$x \rightarrow \text{a random initial configuration } x_{init}^*$

while $T > T_{min}$

$j = \text{a defined maximum number of iterations for the metropolis algorithm } J_{in}$

while $x$ is not a semi-equilibrium configuration AND $j > 0$

$x’ = \text{PERTURB}(x)$

if $\Delta E (H(x’, \lambda^*), H(x, \lambda^*)) < 0$

$x = x’$

else

$x = x’$ with probability (proportional to) $e^{-\Delta E/T}$

end if

$j = j - 1$

if $x$ is a semi-equilibrium configuration

$x_{opt} = x$

end if

end while

if $j > 0$

$T = T - 1$

else

return $x_{opt}$

end if

end while

return $x_{opt}$

end procedure

As already stated, an optimal protocol should also satisfy two safety thresholds on the number of cancer cells (CC). These safety thresholds bound to the in silico observed number of CC when a Chronic protocol is administered. If, at any time $t$, the number of CC is greater than any of the safety thresholds the survival time of that mouse $i$ is $\tau_i = t$.

In previous in silico GA experiments, a random sample of $k=8$ mice was selected from a population of 200 virtual mice. To compare the results with those obtained using GA, the SA algorithm was tested on the same 8 random selected virtual mice sample used by GA. The protocol was then validated on the same population set (that is, 200 virtual mice).

Fig. 4 shows a sketch on how a physical observed phenomenon (the annealing) gave the idea of an optimization algorithm and how this methodology can be applied to a biological problem.

As before, in the GA experiments, the in silico tumor-free percentage of individuals was of 87%. The SA in silico protocol obtained similar results (86.5%) showing almost no difference with GA results. The mean number of cancer cells was computed on the 200-mice set for the GA-protocol and the SA-protocol. Only the SA-protocol demonstrated able to fulfill the safety threshold conditions.

Moreover the SA algorithm required about 2 h on a 8 processor unit to find a protocol with 37 vaccine administrations. Comparison with previous GA results shows a speed-up factor of $\approx 1.4 \cdot 10^2$ (Pennisi et al., 2008).

5. Conclusions

Optimization theory has a long tradition and the techniques are numerous. Most of the practical problems in physics, engineers and applied mathematicians can be formulated as optimization problems. From this perspective the search for an optimized therapeutic protocol for the administration of a vaccine is no exception. In this article we have tried to show how this search for the best vaccine administration in terms of dosage and timings can be formulated as an optimization problem and then, how it can be solved using well-known mathematical or computational methods.

Vaccines are devised to induce an immune system response, thus the finding of an optimal vaccination protocol requires the construction of a mathematical and/or computational model of the immune response. In this article we have described examples of both the discrete approach (i.e., models based on agents) and the continuous approach (i.e., based on ordinary differential equations).

The two modeling approaches are somehow complementary; the continuous ones describe the average behavior of the system and are more akin to a mathematical treatment, while the discrete approaches, like the agent-based, track the individual fate of each biological entities and are therefore closer to reality although the results obtained are more difficult to be interpreted.

With respect to the optimization of therapy administration protocols, the continuous approach requires artificial intelligence methods like the genetic algorithm and the simulated annealing, whereas the discrete one can count on the use of the optimal control techniques for both the definition and resolution of the optimal problem. In both cases the output of the optimization algorithm is a therapeutic protocol candidate for medical screening and validation.

The quality of the optimized protocol is strictly related to the goodness of the mathematical model. In particular the direct or indirect effects of the therapeutic agents on the malignancy need to be carefully taken into account since the optimization algorithms rely on a scoring method that price the solutions on the basis of their effects. Having this said, it appears clear that the more sophisticated and detailed the model is, the higher the chances to obtain an efficacious effect of the optimized therapy once it goes to the test bed.

We like to conclude by stressing the interdisciplinary nature of the experiences described above and by noting that the contribution of life scientists needs to go beyond the only data supply, as it is extremely important in defining the biological scenario and ultimately construct a robust and validated mathematical or computational model that can be combined with the optimization methods here illustrated to become a software system to optimize vaccine protocols. Only through a common effort of life and computer...
scientists it is possible to turn software into a valuable tool in life sciences.

Acknowledgement

This work was partially supported under the EC contract FP6-2004-IST-4, No.028069 (Immunogrid, http://www.immunogrid.eu) and under EC contract FP6-2005-NEST-PATH, No.043241 (Complex-Dis). Computing resources were provided by Consorzio COMETA.

References


Bhattacharya-Chatterjee M, Chatterjee SK, Foon KA. Anti-idiotype antibody vaccine and under EC contract FP6-2005-NEST-PATH, No.043241 (Complex-Dis). Computing resources were provided by Consorzio COMETA.

References


