Computational approaches for translational oncology: Concepts and patents

COMPUTATIONAL APPROACHES FOR TRANSLATIONAL ONCOLOGY: CONCEPTS AND PATENTS

Scianna Marco[1], Munaron Luca[2*]

[1]Researcher of Mathematical Physics
Dept. Mathematical Sciences
Polytechnic of Torino, ITALY
Corso Duca degli Abruzzi 24
10129 Torino, ITALY
phone +390110907558
email: marco.scianna@polito.it

[2]Associate Professor of Physiology
Dept. Life Sciences and Systems Biology
Centre for Nanostructured Interfaces and Surfaces (NIS)
University of Torino, ITALY
Via Accademia Albertina 13
10123 Torino, ITALY
phone +390116704667
email: luca.munaron@unito.it

*corresponding author

Short running title: Computational Approaches For Translational Oncology

Abstract

Background: Cancer is a heterogeneous disease, which is based on an intricate network of processes at different spatiotemporal scales, from the genome to the tissue level. Hence the necessity for the biomedical and pharmaceutical research to work in a multiscale fashion. In this respect, a significant help derives from the collaboration with theoretical sciences. Indeed, mathematical models can provide insights into tumor-related processes and support clinical oncologists in the design of treatment regime, dosage, schedule and toxicity. **Objective and Method:** The main objective of this article is to review the recent computational-based patents which tackle some relevant aspects of tumor treatment. We first analyze a series of patents which concern the purposing or repurposing of anti-tumor compounds. These approaches rely on pharmacokinetics and pharmacodynamics modules, that incorporate data obtained in the different phases of clinical trials. Similar methods are also at the basis of other patents included in this paper, which deal with treatment optimization, in terms of maximizing therapy efficacy while minimizing side effects on the host. A group of patents predicting drug response and tumor evolution by the use of kinetics graphs are commented as well. We finally focus on patents that implement informatics tools to map and screen biological, medical, and pharmaceutical knowledge. **Results and Conclusions:** Despite promising aspects (and an increasing amount of the relative literature), we found few computational-based patents: there is still a significant effort to do for allowing modelling approaches to become an integral component of the pharmaceutical research.

Key words: cancer treatment, database screening, drug purposing, drug repurposing, mathematical models, pharmacokinetics/pharmacodynamics models, systems biology

1. INTRODUCTION

Cancer is still one of the leading causes of death in the world, and major efforts have been undertaken to improve diagnosis and therapy [1,2]. However, recent technical improvements in high-throughput data generation and scientific advances, accompanied by ever increasing investments, have not yet provided the awaited success [3–5]. This is chiefly because cancer is a complex and heterogeneous disease [6], involving an intricate network of mechanisms that span different spatiotemporal scales that include microscopic, mesoscopic, and macroscopic levels. In particular, the microscopic scale refers to processes that occur at the subcellular level, such as DNA synthesis, duplication and gene mutations. The mesoscopic scale refers to single cell-level phenomena, such as pathological adhesive interactions between cancer cells or between them and ECM components, uncontrolled cell duplication and death, as well as altered cell migration and metastasis. Finally, the macroscopic scale corresponds to multicellular collective dynamics, such as tissue invasion by malignant masses.

Such a multitude of intricate and interacting processes represents a challenge for biomedicine and pharmaceutical companies, that could be helped by the collaboration with mathematical and computational sciences. Indeed, *in silico* simulations, that reproduce the actual biological events, are able to accurately predict measurable features, to reduce the experimental burden and to enable rapid, economical and targeted analyses. Further, mathematical tools, based on testable hypotheses, are inherently quantitative, integrate potentially high dimensional data from multiple sources and allow a mechanistic analysis of the system (be it a pathway snapshot, a cell, an organ or, albeit with limited scope, a virtual human framework). Finally, mathematical models are particularly useful when experimental solutions are either impossible, too complicated to perform, or too expensive. *In silico* approaches can complement or replace experiments when they are dangerous or unethical as well. Multilevel mathematical models can also concur to the optimization of clinical protocols and provide a useful guide for possible future experiments.

The development of a successful cancer model is however a long-term process and requires a standardized and iterative workflow (Fig.1), typical of a multidisciplinary systems biology approach. In more details, the derivation of a powerful mathematical tool first involves the description of experimental observations (*conceptual phase*) and then their translation into formal analytical representations (*formalization phase*). A good approach is to build the simplest model, focused on a single scale/level of abstraction and based on a minimal set of assumptions, with a workable number of independent variables and parameters. As a matter of fact, models characterized by a large enough number of degrees of freedom (free parameters) can in fact be made to fit any data, rendering the entire process more like a curve-fitting exercise [5]. This drawback can be avoided by restricting the complexity of the theoretical method, at least in the initial stages of its development, as simplicity does not necessarily limit performance. After checking its validity and its robustness (in terms of missing data estimation, error and bias reduction) with experimental results (*validation phase*), it is then possible to gradually add more components and more levels of abstraction (*development phase*). Available experimental data can be used not only to guide model design but also to verify and validate computational outcomes. When the established mathematical tool reaches the desired level of details, as well as it is able to qualitatively and/or quantitatively reproduce experimental data, it can be used to predict the results of new experiments (*predictive phase*). Obviously, such predictions have to be challenged and assessed by further experimental analysis: indeed, the computational model should be continuously tested by performing systematic experimental validation. This continuous feedback and feedforward between *in silico* and experimental techniques helps to refine previous knowledge or to generate new hypotheses.

The results of the above-described multidisciplinary workflow depend on the experimental data used in the different steps. In principle, the growing quantities of biological and biomedical outcomes obtained from Omics technologies (high-throughput technologies for characterizing genome, epigenome, transcriptome, proteome, metabolome, and interactome) and from clinical trials are not necessarily matched by an increment in the quality of the resulting information, being data potentially subjected to noise, various design and selection biases [5,7,8]. A preliminary step in a model development should be a quality control over the data used to set up the mathematical approach. For instance, some bioinformatics techniques have been recently established to prevent data overfitting and to deal with filtering and error elimination [9,10].

Despite such issues, systems biology approaches are producing a great amount of knowledge in the context of translational (from bench to bedside) and clinical oncology. In particular, in the last couple of decades, mathematical models have provided novel insights into:

i) descriptive and predictive aspects of tumor growth, progression, angiogenesis and relative immune response [11–24];

ii) target identification and drug discover, with particular emphasis on the prediction of drug effects in humans for speeding up clinical trials [25–27];

iii) development and efficacy of oncological treatment protocols, focusing on the design of treatment regime (including surgery, radiation and chemotherapy and their possible combination) as well as of dosage, schedule, toxicity, sensitivity and resistance to drugs [28–31].

It is useful to notice that the US FDA Critical Path initiative identified *in silico* modeling (application of mathematics, statistics and computational analysis) as a key challenging area in the development of new or improved drugs [32,33].

The increasing interest in biomathematics and the huge related literature are however accompanied so far by a very limited number of theoretical-based patents, which will be discussed in the next section (see also Table 1 for the most relevant).

Table 1. List Of The Most Relevant Patents Reviewed In This Work, Their Focus And Mathematical Ingredients.

Patent reference	Author(s)	Addressed problem	Mathematical ingredients
[34,35]	Zoref et al. Arakelyan et al.	optimization of drug development	PK/PD models
[36]	Agur et al.	optimization of chemotherapy schedule	PK/PD models ODEs for cell proliferation/death dynamics
[37]	Vainas et al.	optimization of docetaxel and G-CSF regimens in metastatic breast cancer (MBC) patients	PK/PD models granulopoiesis model
[38]	Kuhn et al.	optimization of anti-cancer drug combination	kinetic models
[39]	Hopkins et al.	optimization of pharmaceutical investigation	map of knowledge

2. MATHEMATICAL-BASED PATENTS

Drug development is a lengthy and expensive process. In particular, most of the costs and time needed to develop successful anti-tumor drugs are spent in clinical trials whose failure rate exceeds 80% [40]. The conventional method for clinical trials design actually suffers some relevant drawbacks. On the one hand, a “trial and error” approach alone can not guarantee that the selected schedules are better than other treatment regimens yet to be tried. Furthermore, the number of compounds which can be empirically tested is negligible compared to the potential spectrum of powerful ones. Finally, the “trial and error” method does not allow drug repurposing, the determination of a novel use for an existing drug or the rescuing of a failed compound.

Agur and coworkers have proposed a series of patents dealing with a computational-based interactive technique for purposing a new compound and for repurposing a drug, with the aim of optimizing drug development from the pre-clinical phases through Phase-IV [41,34,35]. In particular, purposing of a drug involves specific clinical indications and/or patient populations along with a treatment schedule. The repurposing of a compound instead involves a change either in the current application of an approved drug or in the intended/tested application of a drug under development. The disclosed techniques are based on more than two decades of biomathematical research in the area of disease control optimization and drug modeling: they are finalized to simulate the dynamics of key biological, pathological and pharmacological processes in a patient undergoing drug treatment [42–51]. The first stage of the mathematical-based approach consists of creating a model of drug pharmacokinetics (PK) and pharmacodynamics (PD). In particular, the PD module is designed starting from putative mechanisms of drug activity retrieved from *in vivo* and *in vitro* studies. The parameters of drug effect on the different target tissues and tumors are empirically estimated and introduced into the model, which already gives a first estimate of the efficacy of the treatment, thereby interactively guiding the empirical research to reveal further data. Then, animal studies can be used for setting up the PK module and to adjust the PD component. Data on different tumor types, on the effects of multiple doses and of combinations with other drugs can be implemented as well. A toxicity module is also designed to account for side effects observed during the animal studies. At this stage, the model has already the capacity to make some approximate predictions on drug administration to humans. In particular, known inter-species differences in tissue features are taken into account when simulating the PD/PK model in human domains, in order to consider feasible dose ranges. In particular, the authors claim that such procedure offers an improvement of the traditional LD10/10 initial dose for Phase-I trials, which is often too low to exert any effect on the disease. Further, at this point, the model can be used to predict drug failure due to too toxic doses. Obviously, during the dose-effect testing, the computer model is continuously validated and fine-tuned by the observed results. By the end of Phase-I, a fully verified *in vivo* human model is indeed available, integrating all the existing data in

the PK/PD of the drug: it can also yield reasonable short-term predictions concerning the effects of definite drug administration schedules on disease progression. This involves an exhaustive search in the protocol space (within all the possible treatment schedules) with the aim of searching the therapeutic strategies expected to provide the highest response and the lowest toxicity (or a combination of both or any other desired criteria defined by the user) for the cancer type of interest. At the onset of Phase-II trials, the treatment schedules proposed by the mathematical approach can be applied in short pilot trials, which involve a relatively small number of patients. After the first results are obtained (about 6 months on average), the model has to be adjusted by implementing the new data. Finally, a set of intensive computational realizations is carried out to simulate the disease progression during an extended period of up to two years: their goal is to predict which of the schedules tested in short-term trials are expected to yield the best results in the long-run. At this stage, the effect of each selected schedule is compared with the outcome of the existing therapies for the same indications: in the absence of benefits, a "NoGo decision" is recommended. The remaining treatments are instead further tested in Phase-III. Compared with the "trial and error" design, the implementation of the proposed mathematical-based interactive clinical approach is expected to save both time and number of patients engaged. Additional advantages derive from the use of Bayesian statistics, that integrate all the available biological, medical, pharmacological, and clinical information in a predictive perspective [42].

As gene therapy still faces significant hurdles, cancer therapy greatly depends on chemotherapeutic methods, that suffer of poor selectivity and severe side-effects. Special drugs are indeed designed to alleviate some of the related symptoms. The already cited Agur's group has developed a patent [52] aimed to optimize chemotherapy schedule by a protocol that minimizes the cytotoxic effects on normal cells in the case of specified conditions and limitations. The technique is based on a mathematical model and the relative computer-based software. In particular, two generic types of cells are considered: the host cells and the target (tumor) cells. Both populations are further subdivided into cycling and resting (quiescent) cells. Each cell can switch phenotype according to its chronological age. Indeed, proper ordinary differential equations (ODEs) allow to calculate the number of cells in each subcompartment at every time interval starting from an initial distribution. A probability vector accounts for the variability of cycle lengths while retaining a deterministic approach. Tumor heterogeneity is included by differentiating malignant individuals into subpopulations, each of them characterized by distinct parameters for their state (e.g., oxygenation). In order to simulate cancer treatment, a pharmacologic component is added to the model: in particular, cell-cycle selective and non-selective drugs are taken into account. The distribution of drugs within and around the tumor are modeled by pharmacokinetics/pharmacodynamics approaches, which reproduce drug administration protocols. The effects of chemotherapy on cells are modeled as it follows: cell-cycle unspecific drugs affect all cell types (normal, malignant, quiescent, proliferating) by blocking their cycle at different stages and causing their death (with different delays). On the other hand, cell-cycle specific drugs kill only proliferating cells, both normal and altered. Additional types of drugs or events relevant for drug kinetics and dynamics (e.g., rate of absorption, development of tumor resistance, etc.) can be introduced into the model. In most cases of anticancer chemotherapy, the dose-limiting toxicity relates to bone marrow suppression: indeed granulopoiesis and thrombopoiesis can be chosen as an example of host cells, in order to predict the negative effect of the treatment on them. Model parameters can be estimated from experimental studies performed in patient populations. However, due to the huge degree of heterogeneity between malignant tumors, even among similar types, and between patients, it would be advantageous to adjust the treatment protocol to the individual case, employing patient-specific data such as age, weight, gender, previous reaction to treatment, molecular and genetic markers. The computerized method is indeed set up to check any given treatment and to choose a very good one according to user's criteria. The goal is achieved through an operation research tool, the fitness function, which defines how the different protocols score to predefined objectives, including the minimal total amount of drug needed for tumor treatment, the shortest period of disease, the smallest toxicity, the minimal the damage to the BM cells.

A further side effect of chemotherapy (and of radiotherapy as well) is neutropenia, that enhances the susceptibility to microbial infections. In particular, neutropenia is the dose-limiting toxicity of the tri-weekly docetaxel (Taxotere®) schedule. A common neutropenia alleviating therapy is based on the use of Granulocyte Colony Stimulating Factor (G-CSF), mainly administered one day post-docetaxel, for 5-6 consecutive days [53,54]. In this context, Agur and colleagues patented a computational method for predicting docetaxel-induced neutropenia, and used the model to identify improved docetaxel and G-CSF regimens in metastatic breast cancer (MBC) patients [37], see also [55,56]. Their computerized framework integrates and interfaces pharmacokinetics/pharmacodynamics (PK/PD) approaches with a mathematical granulopoiesis model. In more details, a three-compartment PK/PD module [57,58] is designed to calculate the concentration profile of each compound, with the central and the two peripheral compartments respectively representing blood vessels and all body tissues that have direct and fast exchange with blood, such as the BM. In particular, equal or proportional

compartmental concentrations can be assumed, whereas drug distribution is described as a linear exchange between the connected domains. Following the related literature (see, for instance, [59–63] and references therein), the granulopoiesis model reproduces the complex dynamics of mitotic and non-mitotic progenitors and of blood neutrophils, with explicit terms of cell-cycle phases. G-CSF is finally modeled as a feedback molecule governing BM maintenance of steady neutrophils level in blood, accounting for its secretion, diffusion, clearance and interaction with different cell compartments. The cell-cycle structures incorporate experimental data on cell-cycle phase distribution in different BM compartments as well as results from radioactive labelling experiments. The granulopoiesis model parameters are instead estimated by a curve-fitting process based on a variety of *in vivo* experiments, as specified in [43]. The model prediction accuracy was originally tested by dividing a set of patients (mainly Caucasian females) into a "training set", whose clinical outcomes were used to adjust model parameters, and a "validation set", whose medical data were compared to the model-predicted neutrophil profiles upon simulation of docetaxel treatment [64]. Summing up, the input data needed to the computational tool comprise only the patient baseline neutrophil count and the ascribed docetaxel/G-CSF schedule, whereas the output of the software is the effect of the therapy of interest.

Multi-drug chemotherapy has been shown to be a more effective treatment than single drug-based ones: indeed, research efforts have endeavored to develop methods and models for selecting optimal drug combinations, doses, and schedules. Patent [65] describes a computational approach for tailoring optimal cancer multi-drug therapy regimens to individual patients. It can be considered a kinetically tailored treatment ("KITT") model, which predicts therapeutic schedules with the potential to reduce tumor size and prolonging the patient's life, while minimizing side-effects. The disclosed technique is also expected to reduce the number of clinical trials to be performed and to produce robust predictions before *in vivo* studies. The invention includes at least three types of anti-cancer drugs: cell-cycle phase-specific (CS), cell-cycle non-specific (nCS), and cytostatic (DR). Pharmacokinetics and pharmacodynamics models can be then set up for each compound of interest employing empirical measurements taken directly from patients. In particular, the drug concentration is calculated according to the procedure described in [66]. Tumor cells are grouped into two compartments, proliferating and quiescent. Proliferating cells may divide, enter the quiescent state, or die from apoptosis, while quiescent cells may die from necrosis or shift to the proliferating state. In addition, nCS drugs kill cells both in the resting and in the proliferating states, whereas CS drugs kill only a fraction of proliferating cells in a particular phase of their cell cycle. Finally, DR drugs slow cell progression. Tumor individuals are further classified according to their genetic make-up, which determines their resistance level to each type of drug. Resistance is assumed to gradually evolve following a probabilistic law. However, random mutations may increase the level of cell resistance to a given compound. From a computational viewpoint, chemotherapy is set to begin when a tumor reaches a size of 10^9 cells. Cure success is a stochastic event that depends on the minimum tumor size (the nadir) achieved during treatment. In particular, the probability of cure is given by the zeroth term of the Poisson distribution, i.e., $\Pr(\text{cure})=\exp(-\text{nadir})$. This probability declines very quickly as the nadir increases above one cell, and is less than 10^{-4} for a nadir of only 10 cells. The model measures toxicity in terms of dose intensity, plasma drug concentrations, and nadir of the bone marrow-stem cell number. The hematopoietic cell population is assumed to have the same kinetics as the tumor cell populations although with different parameters. Surgical tumor resection is instead simulated as the removal of the oldest 99% of tumor cells, leaving 1% survival of the most recently produced individuals, which are then exposed to adjuvant chemotherapy. The youngest cells are assumed to survive both because they can likely metastasize before the extraction of the older bulk of the tumor and because they arise after many generations of clonal evolution, thereby being more resistant than their progenitors. Finally, the model predicts uncured tumor growth until a lethal size of 10^{12} cells. Then, a computer program product solves the overall mathematical model, needing only patient-specific proliferative/apoptotic rates, cell cycle parameters and levels of drug resistance as required inputs, all of which are data routinely requested by clinical oncologists. The KITT framework is able to analyze a plurality of treatment regimens, each having a quantitative efficacy value associated therewith, as the probability of survival, or survival duration, and the minimization of side effects. Further, the proposed technique can aid in predicting how rapidly resistance to a particular drug may evolve and in studying potential ramifications of treatment modification in a given patient. However, the patent is not free of some shortcomings. First, a spatial component is not present in the KITT model embodiment, therefore neglecting drug diffusion and cell movement. Secondly, no biochemistry is included in the mathematical framework so that, for example, biochemical modulation and synergism between drugs are not incorporated. Finally, the KITT model does not describe host response to chemotherapy in terms of immunological reaction.

Angiogenesis, or neovascularization, is the process of new blood vessels formation, typically promoted by the redundant activity of several growth factors, including vascular endothelial growth factor (VEGF), a potent inducer of endothelial cell proliferation and migration, preferentially expressed in hypoxic areas. The newly formed capillaries then undergo maturation, that involves pericyte coverage and stabilization. The major pericyte-stimulating factor is the platelet-derived growth factor (PDGF). Interactions between endothelial cells and pericytes are also governed by the angiopoietin system composed by two

soluble factors, Angiopoietins 1 and 2 (Ang1 and Ang2, respectively). Ang1 promotes vessel maturation, while Ang2 acts as its natural antagonist. Angiogenesis occurs both in physiological events and in diseases. On the one hand, it allows survival of normal tissues when they become ischemic: on the other hand, pathological vascularization enables tumor growth and metastasis [67,68]. Beside their limitations, apparent advantages of approaches targeted to tumor angiogenesis include their universality for different solid tumors and the lack of prominent side effects and of resistance during repetitive treatment cycles. In this respect, the disclosed teachings in [36] provide a computer-implemented method for determining an optimal treatment protocol of a disease related to angiogenesis, including anti-angiogenic cancer therapies. Such a theoretical approach is based three interconnected modules, which respectively describe tissue growth, angiogenesis (immature vessels growth) and maturation (formation and stabilization of mature vessels). The tissue module simulates tissue cell proliferation and death, which are time- and nutrient-dependent, cell-type specific, and genetically determined. In particular, cell duplication and apoptosis rates are directly and inversely proportional, respectively, to the effective vascular density (EVD), which is a measure of the perfused part of the vascular tree. In the tissue module, VEGF and PDGF production are inversely related to EVD, so that increasing nutrient depletion results in an enhanced secretion of pro-angiogenic factors and consequent enhancement of vessel remodeling and maturation. The angiogenesis module deals with immature vessel volume, which is set to increase proportionally to VEGF concentration (if VEGF is above a given threshold level) and to decay if VEGF is below a possibly different threshold level (generally referred to as “survival level”). In the maturation module, it is assumed that maturation of immature vessels occurs if both pericytes concentration and Ang1/Ang2 ratio are above their respective threshold levels: in this considition, immature vessels do not undergo maturation, while mature vessels are destabilized and become immature. Volume of mature vessels is indeed coherently calculated according both to a given pericyte concentration and to the Ang1/Ang2 ratio. In particular, pericytes proliferate proportionally to PDGF concentration whereas Ang1 and Ang2 are secreted both by nutrient-depleted tissue cells and by immature vessels. Globally, the inputs needed by the computational method are the initial values of tissue volume and of blood vessels density in addition to selected chemical parameters. The outputs are the time evolution of mature and immature vessels sizes and the EVD. The efficacy of pro- or anti-angiogenic drugs (anti-VEGF signaling compounds) can be reproduced by setting their schedule (doses, number of concentrations, and treatment interval) and their effect on the proper mechanisms included in the model (inhibition of VEGF expression) as well as by analyzing the variation of the model outcomes compared to the no-treatment case.

The prediction of drug response in individual patients can be improved by the integration of next generation sequencing (NGS) techniques, which allow large-scale analysis of tumor/patient genomes, epigenomes and transcriptomes. This issue has been addressed by a wide range of different strategies. An approach consists of a simple correlation of one or few biomarkers with published treatment outcomes or with mutational or transcriptomic profiles. Another one involves pattern matching or machine learning algorithms to find optimal drug treatment according to matching transcriptome or genome profiles. However, these procedures display unavoidable limitations, since combinations of biomarkers are either highly correlated (thereby allowing to subdivide a patient population in two or few groups) or not (thereby allowing the clusterization into too small groups, from a statistical viewpoint). More severely, such approaches do not take into account complex data on regulation and connectivity of cancer pathways. A patent by Kuhn and colleagues [38] suggests a computer model able to identify a therapeutic drug combination against a tumor, even if the disease includes alterations (mutations, overexpression, fusions, epigenetic changes and/or insertions) in at least two different but crosstalking signaling pathways. In particular, the inventors claim that treatment strategies targeting several signaling pathways in parallel might provide an improved treatment scheme compared with single-drug therapies. In fact, tumors often escape monotherapies due to additional mutations in another pathway, which may redirect the signaling cascade, thereby rendering the effect of an single drug almost ineffective. The present method consists of a kinetic model for the tumor of interest, which is based on a network topology approach. The nodes of the resulting graph represent biological entities, such genes, transcripts, nucleic acids, peptides, proteins, small molecules, complexes, or metabolites. The edges of the network describe the interactions between the biological entities. Interactions include conversion of one or more given biological entities into different ones, possibly under the influence of further biological elements, changes of biological entities amount, for example as a consequence of the action, presence or absence of other biological entities. Interactions between network variables are then modeled using mass action kinetics, which require the knowledge of kinetic laws and constants and of the (starting) concentrations of substrates or reactants. However, the exact values of specific parameters can not be often directly measured. This problem is overcome by a Monte Carlo-based approach, in which the unknown data are drawn from probability distributions, generating random parameter vectors eventually used to model patient states (different treatments or different treatment combinations). The tumor alterations or the effects of a drug therapy are modeled by proper variations in kinetic laws, constants and/or biomolecule concentrations, as known from the literature or by using inferences from bioinformatics technologies. For instance, missense mutations that damage known functional domains can be modeled by removing the appropriate edge between the modeled biological entity, whereas a drug that changes

the efficiency of an enzymatic biological entity is modeled by multiplying the corresponding kinetic constant by a factor experimentally determined or selected from a lognormal distribution. Also in this case, an optimization procedure is finally employed to establish the more promising drug schedules. It is possible to simulate the effect of the combination of at least two chemotherapeutic compounds in the overall tumor network and to analyze how the multi-drug treatment affects the signaling pathway crosstalk as well. The derivation and the subsequent use of kinetic models of biological networks for clinical purposes is addressed also in patent [69]. Interestingly, the authors claim that, by the use of their approach, the construction of a "kinetic graph" of a biological system can be done even if only one part of the kinetic constants and/or of the starting concentrations are known or derived from experimental data. By employing a Monte Carlo method, the remaining fraction of parameters can indeed be chosen from appropriate probability distributions, including uniform, exponential, Poisson, Binomial, Cauchy, Beta, Gaussian, and lognormal laws. Indeed, the usual outcomes of the biological network, such as steady-state or equilibrium concentrations of selected biological entities, depend on the values assigned to the unknown, randomly chosen, parameters. The unknown data can be optimized by minimizing the difference between predicted and observed quantities. This process may involve continuous optimization and therefore non-linear regression-type approaches. Alternatively or in addition, optimization may involve discontinuous steps, such as modifications of the topology of the network and/or of kinetic laws. Therefore, the proposed patent can be used to evaluate unknown effects of drugs on selected biological entities. Further, this invention allows the simulation of interactions between networks, which may represent different cells, leading to a description of multi-cellular assemblies, tissues, organs, entire organisms or populations.

Drug repositioning is the issue tackled by patent [70]. In particular, the invention allows to define drugs or drug combinations for treating unmet medical needs. The proposed method first involves the selection of the diseases to be treated, including cancer, neurodegenerative disorders, and neuropathies. Then, a dynamic model of the disease is built to define the most relevant mechanisms or cell targets by which the disease may be influenced or corrected by therapy. In particular, it is possible to focus on disease-associated pathways or groups of pathways rather than on isolated gene(s). The proposed technique allows to employ a global statistical analysis of all SNPs (single nucleotide polymorphisms) of each pathway accounted in the model of the disease. Other methods of gene product identification such as proteomic, metabolomic etc., are equally suitable for the identification of the disease-related pathways as well. An *in silico* screening of drugs affecting the pathways, identified as relevant for therapeutic interventions, can be then accomplished. This test can be performed from a library of approved drugs by making hypotheses on their use, known activities and/or putative synergic effects. Finally, the candidate compounds are tested in suitable *in vitro* and/or *in vivo* disease models, in order to evaluate their potential efficacy and toxicity.

Some attempts to improve the efficiency of the drug discovery and development procedure also employ large-scale computing technology, in particular bioinformatics tools and infrastructures. The authors of the patent [39] report a method for a computer-assisted biomedical investigation that consists in mapping pharmaceutical knowledge into a three-dimensional domain, whose coordinate axes pertain disease, targets, and drug compounds, respectively. In order to achieve a complete and systematic analysis, the axes themselves are set up to be substantially comprehensive. For example, the disease axis may be derived from one or more dictionaries or encyclopedias of pathologies. The compound axis may be obtained from databases of drugs that are being marketed or that have been disclosed as under development. The target axis may be taken from the list of genes and protein products, expressed from one or more genomes, that are known to interact with the compounds on the drug axis. Some embodiments support additional possibilities for the axes, such as anatomy, tissue type, cell type, experimental procedure and so on. Note that the number of entities on an axis can be significantly lower than the number of entities on another axis. An information item is then placed within the three-dimensional space, according to the position on the coordinate axes of the entities it is linked with. Turning this around, the existence of the information item can be the evidence of a correlation between the corresponding entities. Entering in more details, an information item is linked to an entity by performing a textual search in selected literature databases of pharmaceutical, biological and medical research papers. An information item could also be linked to an entity by performing a textual search of one of its synonyms. This mapping is used to integrate various and diverse sources of textual, numerical, and graphical data as well as to specify a candidate hypothesis of the generic formula "A is related to B", where A is selected from an axis and B from a second axis. Therefore, the proposed approach supports the systematic identification of potential indications and other medical utilities for drugs and drug targets, of new combinations of medicaments, of biomarkers and surrogate markers to aid drug discovery, clinical diagnostics and/or patient profiling to identified indication(s). The analysis can address questions of toxicity, adverse effects, or other drug safety data, and/or of drug-drug interactions for identifying adverse effects or undiscovered synergies for the development of therapeutic medicaments as well. A number of filters may be applied to the axes and/or to the search results, in order to improve their usefulness. For example, if the second axis represents targets, those entities for which no compound has been launched are excluded. This focuses

investigations on targets related to available marketed drugs. The search results are generally presented as a list of the values ordered according to the number of information items that support the candidate hypothesis. Results may also be ranked (and/or filtered) using semantic algorithms, which typically generate and rank correlations between terms. Similarly, another patent [71] describes a procedure to create and use knowledge patterns, such as a self-organizing knowledge map, for recognizing previously unseen or unknown correlations within the large amount of pharmaceutical data, obtained by virtual screening. A two-dimensional table that maps compounds against targets is finally proposed in [72], based on a comprehensive screening of experimental results: it can be used, for example, to predict a potential use of a new compound, say "A", since it allows to link the targets to a drug, say "B", that is associated to the compound of interest "A".

Finally, human cancer progression simulation is addressed by patent [73]. In particular, the authors claim that their Human Cancer Virtual Simulation (and related software) predicts tumor growth and metastasis starting from biological and medical research information. In particular, users enter individual patient data and then a series of mathematical algorithms screen selected databases (both genetic, molecular, and statistical) searching for results on the features of the tumor. In this way, information is provided about past, present and the future desease development. The invention has distinct modules, each of them dealing with different tumor aspects: (i) origin (original genetic mutations found to be the initiatory step of the disease); (ii) individual cell properties (cell-cycle timing, rate of mitosis or survival); (iii) colony properties (phenotypic relationships between malignant cells, such as collective nutrient consumption and adhesive forces); (iv-v) properties of the entire tumor and the host tissue (size and shape of the malignant mass and its interactions with the surrounding environment, such as with vascular and lymphatic systems); (vi) metastatic potential (ability to invade the host at distant sites, possible effects of surgery and pharmacological treatments). The underlying assumption (which is also its main drawback) is that cancer behavior is similar to that documented for analogous diseases. The human cancer virtual simulation engine can be continuously refined both by increasing the stored data from the medical/biological research and by improving the mathematical search algorithms.

The occurrence and the clinical outcomes of specific gene expression and mutations in cancer and related diseases is tackled by some of the previously presented patents. Further, it is the focus of other inventions that, strictly speaking, are based on statistical methods, such as Bayesian networks, hidden Markov models and other machine learning techniques, with relative bioinformatics tools. For instance, patent [74] allows to differentiate clinical conditions associated with breast cancer by the use of a computer software for data analysis of microarrays. In particular, the disclosed technique is able to distinguish between BRCA1 and sporadic tumors and to provide information on estrogen receptor ESR1 expression and on the likelihood of malignant distant metastases within five years of initial diagnosis. Invention [75] provides methods for detecting the presence of gene mutations associated to early colorectal cancer. Statistical approaches, including sequential probability ratio testing (SPRT), are also used to provide relevant genetic information on an individual (mainly related to cancer-associated diseases) [76], to evaluate the possibility that progeny of two subjects will exhibit one or more phenotypic attributes, and to determine (by defining a proper cutoff value) whether a nucleic acid sequence imbalance exists within a biological sample [77]. A penalized discriminant analysis, an extension of the Fisher's linear discriminant analysis, is then proposed in [78] for the study of gene expression profile microarrays, in order to define the presence or the stage of a selected disease, disorder, or genetic pathology in a mammalian subject. Such a technique includes the examination of a sample containing patients immune cells and the evaluation of the expression variability in a statistically significant number of genes compared either with average healthy or altered gene expression profiles. Further, in [79], the authors propose a method to determine whether an individual is at risk for developing prostate cancer at a later date or whether he suffers from prostate cancer as a result of PG1 gene mutation. This approach involves a statistical analysis based on the LOD score of chromosomal regions potentially harboring a candidate gene associated with a sporadic trait. Finally, a series of patents by Schadt and coworkers reports statistical methods, mainly multivariate analysis, and relative computer systems to associate a query QTL or a query gene with traits using cross species data [80] and to identify genes and/or pathways associated with a trait [81].

CURRENT & FUTURE DEVELOPMENTS

Theoretical methods can be viewed as a useful tool for translational medicine to identify novel therapeutic targets, to provide insight in the mechanisms underlying drug resistance and to predict alternative therapeutic strategies. From a computational point of view, anti-cancer drug development is an optimization problem which involves the maximization of cancer cell death and concomitantly the minimization of toxicity levels. However, despite the aforementioned promising beginnings, there is a low number of patents involving mathematical approaches in drug discovery. A possible explanation is that the outcomes of purely modelling approaches are not yet sufficiently accurate and/or reproducible in experimental trials. Further, the oversimplification of complex biological events, which is often necessary due to computational limitations, gives rise to implausible or unfeasible

evidence. Finally, theoretical methods are usually based on a wide range of parameters and coefficients lacking direct biological correspondence, being therefore difficult to be properly and realistically estimated. For these reasons, mathematical-based approaches still need more time and effort to be more convincing and to become an integral component of the pharmaceutical practice. A significant improvement will be achieved through a closer collaboration between researchers from different backgrounds, leading to multidisciplinary work teams.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest

ACKNOWLEDGEMENTS

The authors thank Prof. Luigi Preziosi (full professor at the Department of Mathematical Sciences of the Polytechnic of Torino) for the useful discussion on the mathematical aspects of the review.

REFERENCES

[1] Gallasch R, Efremova M, Charoentong P, Hackl H, Trajanoski Z. Mathematical models for translational and clinical oncology. J Clin Bioinforma 2013;3:23. doi:10.1186/2043-9113-3-23.

[2] Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. CA Cancer J Clin 2012;62:10–29. doi:10.3322/caac.20138.

[3] Cartwright ME, Cohen S, Fleishaker JC, Madani S, McLeod JF, Musser B, et al. Proof of concept: a PhRMA position paper with recommendations for best practice. Clin Pharmacol Ther 2010;87:278–85. doi:10.1038/clpt.2009.286.

[4] Whitmore E. Development of FDA-Regulated Medical Products: A Translational Approach. ASQ Quality Press; 2012.

[5] Wang Z, Deisboeck TS. Mathematical modeling in cancer drug discovery. Drug Discov Today 2014;19:145–50. doi:10.1016/j.drudis.2013.06.015.

[6] Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. Cell 2011;144:646–74. doi:10.1016/j.cell.2011.02.013.

[7] Costa FF. Big data in biomedicine. Drug Discov Today 2014;19:433–40. doi:10.1016/j.drudis.2013.10.012.

[8] Howe D, Costanzo M, Fey P, Gojobori T, Hannick L, Hide W, et al. Big data: The future of biocuration. Nature 2008;455:47–50. doi:10.1038/455047a.

[9] Sokolov A, Funk C, Graim K, Verspoor K, Ben-Hur A. Combining heterogeneous data sources for accurate functional annotation of proteins. BMC Bioinformatics 2013;14 Suppl 3:S10. doi:10.1186/1471-2105-14-S3-S10.

[10] Mostafavi S, Morris Q. Fast integration of heterogeneous data sources for predicting gene function with limited annotation. Bioinformatics 2010;26:1759–65. doi:10.1093/bioinformatics/btq262.

[11] Shumate SD, El-Shenawee M. Computational model of ductal carcinoma in situ: The effects of contact inhibition on pattern formation. IEEE Trans Biomed Eng 2009;56:1341–7. doi:10.1109/TBME.2008.2005638.

[12] Pham K, Frieboes HB, Cristini V, Lowengrub J. Predictions of tumour morphological stability and evaluation against experimental observations. J R Soc Interface 2011;8:16–29. doi:10.1098/rsif.2010.0194.

[13] Deisboeck TS, Wang Z, Macklin P, Cristini V. Multiscale cancer modeling. vol. 13. 2011. doi:10.1146/annurev-bioeng-071910-124729.

[14] Quaranta V, Weaver AM, Cummings PT, Anderson ARA. Mathematical modeling of cancer: The future of prognosis and treatment. Clin Chim Acta 2005;357:173–9. doi:10.1016/j.cccn.2005.03.023.

[15] Anderson ARA. A hybrid mathematical model of solid tumour invasion: The importance of cell adhesion. Math Med Biol 2005;22:163–86. doi:10.1093/imammb/dqi005.

[16] Araujo RP, McElwain DLS. A history of the study of solid tumour growth: The contribution of mathematical modelling. Bull Math Biol 2004;66:1039–91. doi:10.1016/j.bulm.2003.11.002.

[17] Scianna M, Bell CG, Preziosi L. A review of mathematical models for the formation of vascular networks. J Theor Biol 2013;333:174–209.

[18] Scianna M, Preziosi L. A Hybrid Model Describing Different Morphologies of Tumor Invasion Fronts. Math Model Nat Phenom 2012;7:78–104.

[19] Scianna M, Munaron L, Preziosi L. A multiscale hybrid approach for vasculogenesis and related potential blocking therapies. Prog Biophys Mol Biol 2011;106:450–62. doi:10.1016/j.pbiomolbio.2011.01.004.

[20] Kim DS. Cancer vaccines in the immunotherapy era: Rational approach. Hum Vaccin Immunother 2013;9:2017–8. doi:10.4161/hv.25556.

[21] De Boer RJ, Hogeweg P, Dullens HFJ, De Weger RA, Den Otter W. Macrophage T lymphocyte interactions in the anti-tumor immune response: a mathematical model. J Immunol 1985;134:2748–58.

[22] De Pillis LG, Radunskaya AE, Wiseman CL. A validated mathematical model of cell-mediated immune response to tumor growth. Cancer Res 2005;65:7950–8. doi:10.1158/0008-5472.CAN-05-0564.

[23] Pappalardo F, Brusic V, Castiglione F, Sch??nbach C. Computational and bioinformatics techniques for immunology. Biomed Res Int 2014;2014:1–2. doi:10.1155/2014/263189.

[24] Castiglione F, Toschi F, Bernaschi M, Succi S, Benedetti R, Falini B, et al. Computational modeling of the immune response to tumor antigens. J Theor Biol 2005;237:390–400. doi:10.1016/j.jtbi.2005.04.024.

[25] Kumar N, Hendriks BS, Janes KA, de Graaf D, Lauffenburger DA. Applying computational modeling to drug discovery and development. Drug Discov Today 2006;11:806–11. doi:10.1016/j.drudis.2006.07.010.

[26] Materi W, Wishart DS. Computational systems biology in drug discovery and development: methods and applications. Drug Discov Today 2007;12:295–303. doi:10.1016/j.drudis.2007.02.013.

[27] Arrell DK, Terzic A. Network systems biology for drug discovery. Clin Pharmacol Ther 2010;88:120–5. doi:10.1038/clpt.2010.91.

[28] Palladini A, Nicoletti G, Pappalardo F, Murgo A, Grosso V, Stivani V, et al. In silico modeling and in vivo efficacy of cancer-preventive vaccinations. Cancer Res 2010;70:7755–63. doi:10.1158/0008-5472.CAN-10-0701.

[29] Bozic I, Reiter JG, Allen B, Antal T, Chatterjee K, Shah P, et al. Evolutionary dynamics of cancer in response to targeted combination therapy. Elife 2013;2013:e00747. doi:10.7554/eLife.00747.

[30] Iliadis A, Barbolosi D. Optimizing drug regimens in cancer chemotherapy by an efficacy-toxicity mathematical model. Comput Biomed Res 2000;33:211–26. doi:10.1006/cbmr.2000.1540.

[31] Enderling H, Hahnfeldt P, Hlatky L, Almog N. Systems biology of tumor dormancy: Linking biology and mathematics on multiple scales to improve cancer therapy. Cancer Res., vol. 72, 2012, p. 2172–5. doi:10.1158/0008-5472.CAN-11-3269.

[32] Woodcock J, Woosley R. The FDA Critical Path Initiative and Its Influence on New Drug Development*. Annu Rev Med 2008;59:1–12. doi:10.1146/annurev.med.59.090506.155819.

[33] Zhang P, Brusic V. Mathematical modeling for novel cancer drug discovery and development. Expert Opin Drug Discov 2014;9:1–18. doi:10.1517/17460441.2014.941351.

[34] Arakelyan L, Selitser V, Agur T, Tzoreff TE. Techniques for purposing a new compound and for re-purposing a drug. US20100161301 (2013) n.d.

[35] Arakelyan L, Selitser V, Agur Z. Interactive technique for optimizing drug development from the pre-clinical phases through phase-IV. US 7970550 n.d.

[36] Agur Z, Arakelyan L, Vainstein V. Treatment protocol generation for diseases related to angiogenesis. US20080275684 (2008) n.d.

[37] Vainas O, Vainstein V, Inbar O, Kleiman M, Radel Ben-Av ZA. Improving cancer therapy by docetaxel and granulocyte colony-stimulating factor (g-csf). US20110286960 (2011) n.d.

[38] Kühn A, Lange B, Dreher F, Peycheva S, Lehrach H, Wierling C. Computational approach for identifying a combination of two drugs. WO2015052129A1 (2015) n.d.

[39] Hopkins A, Harland L, Lanfear J, Groom C, Parsons I, Parsons T, et al. System and method for the computer-assisted identification of drugs and indications. US20050060305 A1 n.d.

[40] DiMasi JA, Hansen RW, Grabowski HG. The price of innovation: New estimates of drug development costs. J Health Econ 2003;22:151–85. doi:10.1016/S0167-6296(02)00126-1.

[41] Zoref TE, Agur Z. Techniques for purposing a new compound and for repurposing a drug. WO2009027843A2 (2009) n.d.

[42] Simon R. Bayesian design and analysis of active control clinical trials. Biometrics 1999;55:484–7. doi:10.1111/j.0006-341X.1999.00484.x.

[43] Vainstein V, Ginosar Y, Shoham M, Ranmar DO, Ianovski A, Agur Z. The complex effect of granulocyte colony-stimulating factor on human granulopoiesis analyzed by a new physiologically-based mathematical model. J Theor Biol 2005;234:311–27. doi:10.1016/j.jtbi.2004.11.026.

[44] Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discov 2004;3:673–83. doi:10.1038/nrd1468.

[45] Dodion P, Kenis Y, Staquet M. Phase I trials of single agents in adult solid tumours: preclinical and clinical aspects. Drugs Exp Clin Res 1986;12:23–30.

[46] Fridborg H, Nygren P, Larsson R. Relationship between pharmacokinetic parameters in patients and cytotoxicity in vitro of standard and investigational anticancer drugs. Anticancer Drugs 1995;6:64–9.

[47] Eisenhauer EA, O'dwyer PJ, Christian M, Humphrey JS. Phase I clinical trial design in cancer drug development. J Clin Oncol 2000;18:684–92.

[48] Simon R. Accelerated Titration Designs for Phase I Clinical Trials in Oncology. J Natl Cancer Inst 1997;Vol. 89 No:1138–47. doi:10.1093/jnci/89.15.1138.

[49] Collins JM, Zaharko DS, Chabner BA, Dedrick RL. Potential roles for preclinical pharmacology in phase I clinical trials. Cancer Treat Rep 1986;70:73–80.

[50] Norel R, Agur Z. A model for the adjustment of the mitotic clock by cyclin and MPF levels. Science (80-) 1991;251:1076–8.

[51] Agur Z, Arnon R, Sandak B, Schechter B. Zidovudine toxicity to murine bone marrow may be affected by the exact frequency of drug administration. Exp Hematol 1991;19:364–8.

[52] Agur Z, Fleishmann S, Skomorovski K, Vardi M. System and methods for optimized drug delivery and progression of diseased and normal cells. US20120284005 (2012) n.d.

[53] Paciucci PA, Raptis G, Bleiweiss I, Weltz C, Lehrer D, Gurry R. Neo-adjuvant therapy with dose-dense docetaxel plus short-term filgrastim

rescue for locally advanced breast cancer. Anticancer Drugs 2002;13:791–5. doi:10.1097/00001813-200209000-00002.

[54] Gridelli C, Frontini L, Barletta E, Rossi A, Barzelloni ML, Scognamiglio F, et al. Single agent docetaxel plus granulocyte-colony stimulating factor (G-CSF) in previously treated patients with advanced non small cell lung cancer. A phase II study and review of the literature. Anticancer Res 2000;20:1077–84.

[55] Ferree SM, Cowens JW, Jorgensen CLT, Nielsen TO, Ejlertsen B. Methods of Treating Breast Cancer with Gemcitabine Therapy. US20140037620A1 (2014) n.d.

[56] Perou CM, Bernard PS, Nielsen TO, Ellis MJ, Parker JS, Martin M, et al. Methods of treating breast cancer with taxane therapy. US 9181588 B2 n.d.

[57] Gorelik B, Ziv I, Shohat R, Wick M, Hankins WD, Sidransky D, et al. Efficacy of weekly docetaxel and bevacizumab in mesenchymal chondrosarcoma: A new theranostic method combining xenografted biopsies with a mathematical model. Cancer Res 2008;68:9033–40. doi:10.1158/0008-5472.CAN-08-1723.

[58] Brunsvig PFR, Andersen A, Aamdal S, Kristensen V, Olsen H. Pharmacokinetic analysis of two different docetaxel dose levels in patients with non-small cell lung cancer treated with docetaxel as monotherapy or with concurrent radiotherapy. BMC Cancer 2007;7:197. doi:10.1186/1471-2407-7-197.

[59] Schmitz S, Franke H, Brusis J, Wichmann HE. Quantification of the cell kinetic effects of G-CSF using a model of human granulopoiesis. Exp Hematol 1993;21:755–60.

[60] Friberg LE, Henningsson A, Maas H, Nguyen L, Karlsson MO. Model of chemotherapy-induced myelosuppression with parameter consistency across drugs. J Clin Oncol 2002;20:4713–21. doi:10.1200/JCO.2002.02.140.

[61] Engel C, Scholz M, Loeffler M. A computational model of human granulopoiesis to simulate the hematotoxic effects of multicycle polychemotherapy. Blood 2004;104:2323–31. doi:10.1182/blood-2004-01-0306.

[62] Shochat E, Rom-Kedar V. Novel strategies for granulocyte colony-stimulating factor treatment of severe prolonged neutropenia suggested by mathematical modeling. Clin Cancer Res 2008;14:6354–63. doi:10.1158/1078-0432.CCR-08-0807.

[63] Ostby I, Rusten LS, Kvalheim G, Grottum P. A mathematical model for reconstitution of granulopoiesis after high dose chemotherapy with autologous stem cell transplantation. J Math Biol 2003;47:101–36. doi:10.1007/s00285-003-0198-6.

[64] Bonate PL. Pharmacokinetic-Pharmacodynamic Modeling and Simulation. Springer; 2011.

[65] Gardner SN. Computational model, method, and system for kinetically-tailoring multi-drug chemotherapy for individuals. US7286970B2 (2007) n.d.

[66] Hryniuk WM. Average relative dose intensity and the impact on design of. SeminOncol 1987;14:65–74.

[67] Jain RK. Antiangiogenesis Strategies Revisited: From Starving Tumors to Alleviating Hypoxia. Cancer Cell 2014;26:605–22. doi:10.1016/j.ccell.2014.10.006.

[68] Jain RK, Carmeliet P. SnapShot: Tumor angiogenesis. Cell 2012;149:1408–1408.e1. doi:10.1016/j.cell.2012.05.025.

[69] Lehrach H, Herwig R, Wierling C. Computer implemented model of biological networks. WO2010025961 A2 n.d.

[70] Cohen D, Chumakov I. Novel disease treatment by predicting drug association. WO2009068659A3 (2009) n.d.

[71] Mark A D Collins, John R Shaw RLC. Method and system for creating and using knowledge patterns. US6768982 (2002) n.d.

[72] Hao Chen DM. Identification of molecular targets useful in treating substance abuse and addiction. US20020187514A1 (2002) n.d.

[73] Meagher JF, Thomas AW, Thomas J, Thomas RD, Thomas S. Human cancer virtual simulation system. WO2001000083 A1 n.d.

[74] Dai H, He Y, Linsley PS, Mao M, Roberts CJ, Van't VLJ, et al. Diagnosis and prognosis of breast cancer patients. WO 200210332A2 n.d.

[75] Lapidus SN, Shuber AP, Ulmer KM. Method for the detection of clonal populations of transformed cells in a genomically heterogeneous cellular sample. US5670325A (1997) n.d.

[76] Reese M, White C. Methods of selection, reporting and analysis of genetic markers using borad-based genetic profiling applications. US20070042369A1 (2007) n.d.

[77] Lo YMD, Chiu RWK, Chan KC, Zee BCY, Chong KC. Determining a nucleic acid sequence imbalance. US20090087847A1 (2009) n.d.

[78] Showe L, Showe M, Kari L, Nebozhyn M, Loboda A. Method of diagnosis of cancer based on gene expression profiles in cells. US20060271309A1 (2006) n.d.

[79] Cohen D, Chumakov I, Blumenfeld M, Bougueleret L. Prostate cancer gene. US5945522A (1999) n.d.

[80] Schadt E, Monks S, Lamb J. Computer systems and methods for associating genes with traits using cross species data. US20070166707A1 (2007) n.d.

[81] Schadt EE, Monks SA. Computer systems and methods for identifying genes and determining pathways associated with traits. US7035739B2 (2006) n.d.

Figure legend

Figure 1: The systems biology workflow. *In silico* and *in vitro* approaches are integrated and interfaced in a continuous feedback and feedforward of information. In particular, initial experimental observations, translated in mathematical language, constitute the hypothesis of a first simple model. After multiple cycles of experimental validation, the computational method can be finally used in a predictive way to improve biomedical knowledge.