PROCEEDINGS

of the
2012 5th International Advanced Research Workshop on *In Silico* Oncology and Cancer Investigation – The TUMOR Project Workshop (IARWISOCI)

An IEEE-EMBS Technically Co-sponsored Conference
Funded by the European Commission
through the Transatlantic TUMOR Project
in the Framework of the VPH Initiative

Athens, Greece, 22-23 October 2012

Edited by G. S. Stamatakos and D. D. Dionysiou

(Open Access Version)

Available at www.5th-iarwisoci.iccs.ntua.gr

EMB IEEE

2012 5th International Advanced Research Workshop on *In Silico* Oncology and Cancer Investigation - The TUMOR Project Workshop (IARWISOCI)
(**open-access version**)

Edited by Georgios S.Stamatakos and Dimitra D. Dionysiou

Available at www.5th-iarwisoci.iccs.ntua.gr

ISBN (electronic): 978-618-80348-0-8

Institute of Communication and Computer Systems
National Technical University of Athens
(ICCS-NTUA)
Iroon Polytechniou 9
Zografos
GR-157 80 Greece

Proceedings of the 2012 5th International Advanced Research Workshop on *In Silico* Oncology and Cancer Investigation – The TUMOR Project Workshop (IARWISOCI)

Athens, Greece, 22-23 October 2012

Edited by
G. S. Stamatakos and D. D. Dionysiou
(Open Access Version)

ISBN (electronic): 978-618-80348-0-8

Available at www.5th-iarwisoci.iccs.ntua.gr

An IEEE-EMBS Technically Co-sponsored Conference
Funded by the European Commission
through the Transatlantic TUMOR Project
in the Framework of the
Virtual Physiological Human (VPH) Initiative

ORGANIZING COMMITTEE

General Chair

G. Stamatakos, PhD, ICCS - National Technical University of Athens (GR)

Members

N. Graf, MD, University Hospital of Saarland (DE)

K. Marias, PhD, Foundation for Research and Technology – Hellas (GR)

M. Akay, PhD, University of Houston (USA)

R.Radhakrishnan, PhD, University of Pennsylvania (USA)

D. Dionysiou, PhD, ICCS - National Technical University of Athens (GR)

V. Sakkalis, PhD, Foundation for Research and Technology – Hellas (GR)

N. Uzunoglu, PhD, ICCS - National Technical University of Athens (GR)

PAPER CITATION FORMAT
(Open Access Version)

(Co-)Author(s), "Paper Title." In G. Stamatakos and D. Dionysiou (Eds): Proc. 2012 5th Int. Adv. Res. Workshop on *In Silico* Oncology and Cancer Investigation – The TUMOR Project Workshop (IARWISOCI), Athens, Greece, Oct.22-23, 2012 (www.5th-iarwisoci.iccs.ntua.gr), pp.xx-xx. ISBN: 978-618-80348-0-8 (open-access version)

PUBLICATION NOTE

It is noted that the **open-access version** of the proceedings is freely available on the workshop website www.5th-iarwisoci.iccs.ntua.gr

An **IEEE Xplore® version** of the proceedings containing most of the open-access version content will also be made available.

Please clearly mention which proceedings version a paper citation refers to.

CORRESPONDENCE

All correspondence should be addressed to

Georgios S. Stamatakos
Research Professor
Leader, *In Silico* Oncology Group
Institute of Communication and Computer Systems
National Technical University of Athens
Iroon Polytechniou 9, Zografos GR 157 80 , Greece
Tel: (+30) 210 772 2287, Fax: (+30) 210 772 3557
E-Mail: gestam@central.ntua.gr
URL: www.in-silico-oncology.iccs.ntua.gr

TABLE OF CONTENTS
(Open Access Version)

In Silico Oncology - Leading the Way at the Dawning of *In Silico* Medicine

Editorial

Proceedings of the 2012 5th International Advanced Research Workshop on *In Silico* Oncology and Cancer Investigation – The TUMOR Project Workshop (IARWISOCI)*

Georgios S. Stamatakos, *Member, IEEE*

I. Introduction

Cancer is a *natural phenomenon* and consequently is amenable to mathematical and computational description. Clinically driven complex multi-scale cancer models are capable of producing realistic spatio-temporal and patient-specific simulations of commonly-used clinical interventions such as radio-chemotherapy. Clinical data-processing procedures and computer technologies play an important role in this context. Following clinical adaptation and validation within the framework of clinico-genomic trials, models are expected to advance the prospect of individualized treatment optimization, this being the long term goal of the emergent scientific, technological and medical discipline of *in silico* oncology.

Treatment optimization is to be achieved through experimentation *in silico* i.e. on the computer. Moreover, provision of improved insight into tumor dynamics and optimization of clinical trial design and interpretation constitute short- and mid-term goals of this new domain.

The IEEE-EMBS technically co-sponsored 5th International Advanced Research Workshop on *In Silico* Oncology and Cancer Investigation (5th IARWISOCI) (www.5th-iarwisoci.iccs.ntua.gr), being also the transatlantic TUMOR project workshop (www.tumor-project.eu), proved an excellent opportunity for contributing to the shaping of the discipline. The presented papers deal with modeling of tumor dynamics and response to treatment from the biochemical to the macroscopic level and from basic science to clinics via information technology. They have been contributed (some by invitation) by top international researchers and research groups. The workshop took place in Athens, Greece on 23-24 October 2012.

II. *In Silico* Oncology

In silico oncology could be formally defined as being "…a complex and multiscale combination of sciences, technologies and clinical medicine intending to simulate malignant tumor growth and tumor and normal tissue response to therapeutic modalities at all biomedically meaningful spatio-temporal scales". Its long term goal is to quantitatively understand cancer and related phenomena and optimize therapeutic interventions by performing *in silico* experiments using clinical, imaging, histopathological, molecular and pharmacogenomic data from individual patients. In order to achieve such an ambitious goal translation of cancer models and *oncosimulators* into the clinical trials arena is a *sine qua non* condition.

III. Towards *In Silico* Medicine

In silico oncology serves as an excellent paradigm of the emergent generic domain of *in silico* medicine, which has been designated as one of the key research areas to be supported by the European Commission's upcoming research funding framework, "HORIZON 2020" (http://ec.europa.eu/research/horizon2020/index_en.cfm).

Horizon 2020 is the financial instrument implementing the Innovation Union, a Europe 2020 flagship initiative aimed at securing Europe's global competitiveness. Running from 2014 to 2020 with an €80 billion budget, the EU's new program for research and innovation is part of the drive to create new growth and jobs in Europe. A branch of the program is entitled: "Using *in-silico* medicine for improving disease management and prediction." The current and future importance of both *in silico* oncology and *in silico* medicine is therefore unquestionable.

IV. Organizing Committee

The Organizing Committee of the workshop consisted of the following persons

G. Stamatakos, PhD, ICCS - National Technical University of Athens (GR), General Chair
N. Graf, MD, University Hospital of Saarland (DE)
K. Marias, PhD, Foundation for Research and Technology – Hellas (GR)
M. Akay, PhD, University of Houston (USA)
R.Radhakrishnan, PhD, University of Pennsylvania (USA)
D. Dionysiou, PhD, ICCS - National Technical University of Athens (GR)
V. Sakkalis, PhD, Foundation for Research and Technology – Hellas (GR)
N. Uzunoglu, PhD, ICCS - National Technical University of Athens (GR)

*The workshop was funded by the European Commission through the transatlantic TUMOR project (FP7-IST-247754).

G.S.Stamatakos is with the Institute of Communication and Computer Systems, National Technical University of Athens, 9, Iroon Polytechniou, Zofrafos, 157 80, Greece (phone:+30 210 772 2287; fax: +30 210 772 3557; e-mail: gestam@central.ntua.gr).

Computational Methodology for Mechanistic Profiling of Kinase Domain Mutations in Cancers*

Peter J. Huwe, and Ravi Radhakrishnan

***Abstract*— We present a computational modeling and simulation approach to delineate molecular-level mechanisms of activation of protein receptor tyrosine kinases and describe clinical implications of mutations in the Anaplastic Lymphoma Kinase (ALK) receptor tyrosine kinase in pediatric neuroblastoma. We show here that our results shed molecular-level insight into the various mechanisms governing such transforming mutations at the level of kinase activity and are remarkably consistent with experimental observations. We expect that the current study on ALK with suitable validation will transform our computational approach to enable future predictions of driver oncogenic mutations with low false-positive rates, and can hence serve an important in silico tool toward personalized cancer therapy.**

I. Introduction

Deregulation and mutation of receptor tyrosine kinases (RTKs) have been correlated with cancer almost immediately after their discovery and purification in the early 1980s. The v-erbb oncogene in the avian erythroblastosis virus that was capable of inducing acute leukemia encoded was found to encode a constitutively active form of the homologous ErbB kinase protein [1]. With the increased study upon RTKs, the correlation between deregulation of RTKs and a variety of ailments and particularly in cancer has only grown stronger. Deregulation of RTKs in cancers can occur at several points: (1) increased ligand production through enhanced local autocrine activation; (2) specific gene translocations to produce kinase fusions with altered signaling profiles; (3) RTK overexpression at the cell surface; (4) mutation of the RTK protein to modulate activity; (5) disregulation of phosphatase and endocytosis mechanisms to increase RTK signal propagation. Clinically identified activating RTK kinase domain mutations have been discovered throughout many cancers [1]. The results form Catalog of Somatic Mutations In Cancer (COSMIC), which has a much more thorough listing for all the mutations and all cancers. The oncogenic mutations cluster near the characteristic aspects of kinase activation. The increased kinase activity increases the dependency of tumor upon the RTK and inhibition of the RTK is a viable route for cancer therapeutics. The epidermal growth factor receptor (EGFR) and one of its small molecule inhibitor, gefitinib, is a canonical example of RTKs, cancer and targeted therapeutics [1]. However, examination of the tumors revealed sets of mutations in the EGFR tyrosine kinase domain revealed that only a sub-set of the tumors harboring these EGFR mutations are exceptionally sensitive to inhibition through Gefitinib [1]. There are several other small molecule tyrosine kinase inhibitors (TKIs) as well as antibodies already approved by the FDA and in use in the clinical settings [1]. Hence, it is important to understand RTK activation mechanisms at the molecular level to help design efficacious therapeutics. Computational methodologies offer a powerful, quantitative, and complimentary alternative for the study of intracellular kinase domains, which if utilized correctly can predict effects of mutations on RTK activation mechanisms [1].

Neuroblastoma is an early childhood cancer that arises from neural crest tissue along the sympathetic chain ganglion in the developing autonomic nervous system. Neuroblastoma is the most common extracranial pediatric solid tumor and the most common malignancy diagnosed in the first year of life [2]. Approximately half of the patients diagnosed with the disease are classified as “high-risk” and exhibit a very aggressive phenotype [3]. Even with the use of intensive chemotherapy, radiotherapy, surgery, and bone marrow transplantation, the 5 year survival rate among these high-risk patients remains quite low—a mere ~40% [2,3]. Those fortunate few who do survive the disease suffer chronically from multiple negative sequelae. Currently, no FDA-approved molecularly targeted approaches exist for treating the disease. The need for improved neuroblastoma therapeutics is urgent.

Recently, Maris et al. have identified germline and somatic mutations in neuroblastoma patients that cause increased tyrosine kinase activity in anaplastic lymphoma kinase (ALK) [3]. ALK activation in Neuroblastoma appears to be a result of point mutations rather than gene fusions [2]. Inhibition of ALK in neuroblastoma cell lines carrying amplified or mutated ALK alleles results in compromised downstream signaling and cell growth and ALK has emerged as a potential therapeutic target for neuroblastoma patients.

With a focus on the discovery and exploitation of oncogenic driver mutations, current approaches typically involve collaborative partnership between cancer geneticists, structural, computational and cellular biologists, drug development experts, and clinical trials experts. The different facets of research activities include: (1) Genomics- access to thousands of patient genotypic samples; (2) Patients- large developmental therapeutics team with referrals from multiple nations; (3) Samples- investment in the collection of patient samples including thousands of tumor, DNA, RNA and protein samples, all with rich clinical annotation and outcomes data. This includes a rich bank of dozens of human and murine neuroblastoma-derived cell lines; (4) Clinical trials infrastructure; (5) Structural Biology- structure elucidation of proteins and computational

*Research supported by US NSF/CBET, US NIH/NIBIB, and US XSEDE grants to R. R. and by US NSF/GRF fellowship to P. J. H.

P. J. Huwe is with University of Pennsylvania, Biochemistry and Molecular Biophysics Graduate Group, Philadelphia, PA 19104 (email: pjhuwe@gmail.com)

R. Radhakrishnan is with University of Pennsylvania, Departments of Bioengineering, Chemical and Biomolecular Engineering, Biochemistry and Biophysics, Philadelphia PA 19104 (corresponding author phone: 215-898-0487; fax: 215-573-2071; e-mail: rradhak@seas.upenn.edu).

elucidations of enzymatic mechanisms and studies of signal transduction.

Here, we will focus on the delineation of activation mechanisms using computational techniques and the central problem of oncogenic mutations in RTKs as key mediators of human cancer. Development of computational tools necessary to understand the structural consequences of genetic mutations in cancer will facilitate our quest for molecular-level insight into rational mechanisms for drug development and for combating drug resistance. Current outlook on the receptor tyrosine kinase ALK especially provides an important illustration of how this approach can revolutionize cancer treatment. Namely, oncogenic driver mutations in ALK were identified in neuroblastoma patients, and shown to cause activation of the receptor – immediately indicating the use of existing ALK kinase inhibitors in this patient population. Individual ALK-activating mutations show different inhibitor sensitivities, creating an urgent need to understand the structural basis for their effects in order to guide treatment choice. Emergence of resistance mutations, as also seen with kinase-targeted inhibitors of ABL (in Chronic Mylogenous Leukemia) and EGFR (in Non-Small-Cell Lung Cancer) creates a need to understand their spectrum and the origin of their effects through our integrated approach.

II. Modeling Methods

A. *Molecular Dynamics*

The inactive ALK wild-type tyrosine kinase domain structure was taken from 3L9P PDB (1.80 Å resolution) [4]. Missing loop residues 1117-1122 were modeled on with the loop-modeling algorithm LOOPY, as implemented in MODELLER v9.8. Mutant structures were generated with MODELLER v9.8 by making point mutations to the 3L9P structure. The active ALK homology model was generated with homology modeling program MolIDE and the SCWRL4 rotamer library, using the 1IR3 PDB structure (active insulin receptor kinase) as the template (46% sequence identity; 63% sequence conservation). All structures were modeled in apo form; for details of the methods, see [1,5,6].

All structures were subjected to the same molecular dynamics (MD) protocol. Hydrogen atoms were added to the structures with Automatic PSF Generation Plugin v1.3 implemented in VMD 1.8.6. Consistent with a physiological pH of 7.0, all histidines were modeled with a +1 protonation state on the d-nitrogen. The Solvate Plugin v1.5 and Autoionize Plugin v1.3 implemented in VMD were used to construct an electroneutral waterbox with 15 Å TIP3P water padding and 0.15 M Na^+/Cl^- concentration. All Na^+ and Cl^- ions were placed at least 5 Å away from protein atoms and each other. Systems contained approximately 61000 atoms.

All MD simulations were carried out with NAMD v2.8 molecular dynamics software package using CHARMM27 force field parameters. Periodic boundary conditions were used throughout. The particle mesh Ewald algorithm was used to treat long-range electrostatic interactions. An integration timestep of 2 fs was used. Bonds between hydrogens and heavy atoms were constrained to their equilibrium values, with the velocity correction being performed by the RATTLE algorithm. Rigid waters were treated with the SETTLE algorithm. Long-range nonbonded van der Waals (VDW) interactions were treated by applying a smooth switching function at 10 Å with a cutoff distance of 12 Å. To eliminate unfavorable contacts, the solvated systems underwent an energy minimization using a conjugate gradient algorithm; they were then gradually heated to 300 K. Constant temperature and pressure (NVT) simulations using a Nosé-Hoover Langevin piston were performed at 300 K and 1 atmosphere pressure to equilibrate the volume of the solvation box. Subsequently, constant temperature and volume (NVT) simulations were run on the system. After an equilibration period, 40 ns of NVT simulation were completed on each structure.

B. *Hydrogen Bond Analysis*

Hydrogen bond (H-bond) analysis was performed on the last 1000 frames (20 ns) of trajectory for each system using the HBonds Plugin v1.2 in VMD. Hydrogen-bond cutoff lengths of 3.2 Å and H-O-H angle cutoffs of 150˚ were chosen to include H-bonds of moderate strength. The occupancies (Y) for each residue-to-residue hydrogen bond range from 0% to 100% across the last 20 ns of trajectory in each system. A scoring function was created to analyze how active-like the hydrogen bond networks were for each system. The scoring process is outlined below. (1) For each hydrogen bond, calculate the difference in occupancies between the active (A) wildtype (WT) and inactive (I) WT systems ($\Delta_{WT} = Y^I_{WT} - Y^A_{WT}$). (2) For each bond, if $|\Delta_{WT}| > 40.0\%$, calculate the difference in occupancies between the active WT and inactive mutant (MUT) for each mutation ($\Delta_{MUT} = Y^I_{WT} - Y^A_{MUT}$). (3) If $\Delta_{MUT}/\Delta_{WT} > 0.5$, then the bond receives a binary activation score of 1; otherwise, it receives a score of 0. The scores were tallied separately for the activation loop (A-loop) and the C-helix. Any bonds bridging the two were counted in both analyses. For the A-loop (comprising 15 H-bonds), a collective score of 9 or greater was considered to be activating the kinase. For the C-helix (comprising 7 H-bonds), a collective score of 4 or greater was considered activating. The scores are summarized in Table 1.

C. *Solvent Accessible Surface Area (SASA)*

Solvent accessible surface area (SASA) values were calculated in VMD using the measure SASA module using a probe radius of 1.4 Å. The SASA was calculated on a per-residue basis for the residues forming the hydrophobic core. The SASA values (in units of $Å^2$) were averaged over all steps of the MD trajectory, from which the mean SASA values were computed. Mutant systems were scored as activating as follows: (1) SASA values were normalized using the expression SASA(MUT,I)-SASA(WT,I)/ SASA(WT,A)- SASA(WT,I); (2) If the normalized SASA

scores were less that -2.5 for at least two residues or more, the mutant system was scored as activating (with a score of 1); (3) If the sum of the SASA scores of residues that showed a large difference between the active and inactive states in the WT were negative, the mutant was scored as activating; (4) If the sum in (3) was positive but less than 1.0, but at least one of the residues had a score of <-4.0, the mutant was scored as activating, see Table 1.

III. Categorization and Activation Scores

(1) Activation mechanism based on destabilizing the C-helix interactions in the inactive conformation: in recent computational work on kinase domains of ErbB receptors, we suggested that inactive and active conformations have distinct signatures of the hydrogen bond networks, particularly involving the residues of the C-Helix [5, 6]. A similar analysis was performed on various mutant systems of the ALK kinase domain, see methods, focusing on the C-Helix. A score in this category represents hydrogen-bonding patterns of C-helix residues showing the characteristics of the active-wild type conformation; see Table 1. The mutations were scored p (proximal) if the site of the mutation was part of the C-Helix or proximal (<6 Å) to the C-Helix residues; the choice of 6 Å represents proximity in terms of 1-2 amino acid residues. The mutations in this category are, not surprisingly, mostly clustered proximal to the C-Helix, A-loop, and the active site. However a small number of mutations in the C-lobe and the hydrophobic core are also featured in this category.

(2) Activation based on A-loop interactions: mutations, which were scored in this category, are clustered in regions proximal to the active site and similar to those in the C-Helix category, see Table 1. The mutations were scored p (proximal) if the site of the mutation was part of the A-loop or proximal (<6 Å) to the A-loop residues.

We note that we combine the score based on A-loop and C-Helix interaction using the logical AND function to evaluating the activation status. The rationale for this choice is based on the recognition that A-loop in most kinases is flexible with multiple conformational (locally stable) states, which implies numerous compensatory interactions between active and inactive conformations. Scoring a mutant based on this category based on similarity with the active conformation alone does not take into account the rugged conformational free energy landscape of the A-loop due to the compensatory interactions.

(3) Predictions based on hydrophobic interactions: In our earlier computational work on ErbB kinase systems, we introduced a method to score hydrophobic interactions based on SASA analysis (see methods) [5]. A similar analysis on ALK mutant systems is presented in Table 1. There is a significant difference between the SASA of the residues of the hydrophobic core in the active and inactive conformations. The sum total of the SASA for these residues is significantly higher in the inactive conformation relative to active. Notably, Y1278 (which is one of the tyrosines located on the activation loop) is involved in stacking interactions with some of the residues of the hydrophobic core. In the active conformation, Y1278 is thrust out into solvent causing the hydrophobic core to be significantly more compact. As described in the methods section, mutations that increase the SASA of the hydrophobic core are scored as activating (as they further destabilize the inactive conformation). The mutations were scored p (proximal) if the site of the mutation was proximal (<10 Å) to Y1278; the choice of 10 Å represents the ability to directly impact Y1278 or to disrupt a solvation sphere of that radius surrounding Y1278. The mutations scored in this category are located in the hydrophobic core, but also include a few that are proximal to the active site.

IV. Results and Discussion

We hypothesize that classifying mutations based on similarity in activation scores across different categories (see Table 1) will reveal conserved mechanisms of kinase activation and regulation. Maris et al. [3] have identified >25 mutations to the ALK-tyrosine kinase domain (TKD) in their neuroblastoma patients. Some of the mutants are known to be activating, while the role and significance of others is unclear. Some may only be passenger mutations (i.e. mutations that do not drive the cancer but instead are merely present). Almost all of the activating ALK mutations identified in neuroblastoma patients have been localized to the TKD of ALK. Thus, we focused our computational study on a detailed understanding of activation of the TKD. This approach is important for rational design of inhibitors to block ALK activation in neuroblastoma patients. Because RTKs have dissimilar inactive TKD states, they have different mechanisms of activation. However, some commonalities do exist. Typically, the molecular mechanism in RTK-TKDs for 'switching' from the unique inactive state to the activated state involves autophosphorylation of the activation loop (A-loop) [1]. This autophosphorylation generally disrupts 'cis-autoinhibitory' interactions and also stabilizes the active conformation. We hypothesize that active and inactive states are differentially stabilized by molecular (hydrophilic and hydrophobic) interactions and that activating mutations alter the relative occupancy of these states by destabilizing the inactive or stabilizing the active state. Lemmon and co-workers have shown that the two most common mutants, R1275Q and F1174L, display increased tyrosine kinase activity compared to wild type controls both *in vivo* and *in vitro* [2]. It remains unclear, however, exactly how these mutations functionally cause the kinase to be constitutively active.

Based on the MD simulations of each of the mutant systems (see Table 1) in the inactive conformation, and the WT system in both active and inactive conformations (see modeling methods for details), we have assigned activation scores for the different mutants based on hydrogen bond analysis of activation loop and C-helix residues (see Table 1), based on global motion (data not-shown), and based on hydrophobic analysis (Table 1); these scores enable us to classify the different mutant systems into various sub-

clusters as summarized in Table 1. A score of 1 represents a prediction of activation under the given category. To further qualify the activation scores, we categorize each mutation as proximal (p) or distal (d) for a given category of activation based on the distance between the site of mutation and the sub-region relevant to the category (see Table 1). Based on the activation scores under different categories and based on the proximity of the site of mutation, we record an overall (weighted score) prediction of the activation status (see Table 1). The mutations with activation scores in A-loop and C-Helix categories are scored as activating only if they have recorded an activation score in both categories and are proximal to at least one of the regions (namely A-loop or C-Helix). Mutations with scores in the SASA categories are weighted to yield an overall activation status if they are proximal to Y1278, an important residue impacted by the hydrophobic core. There is a significant difference between the SASA of the residues of the hydrophobic core in the active and inactive conformations. Notably, Y1278 (which is one of the tyrosines located on the A-loop) is involved in stacking interactions with some of the residues of the hydrophobic core. In the active conformation, Y1278 is thrust out into solvent causing the hydrophobic core to be significantly more compact. Our results in Table 1 provide a rational framework not only in terms of categorizing different mutations based on their functional status (namely activating or inactivating), but also categorize them according to a possible mechanism of activation: namely (1) disruption of C-Helix/A-Loop autoinhibitory interactions in the inactive state, (2) destabilizing the hydrophobic core, (3) impacting Y1278 directly independent of hydrophobic effect. Other effect such as (4) impacting nucleotide binding, (5) disrupting the catalytic machinery, (6) distal mutations impacting overall kinase conformation are not considered in the current study and will be pursued in the future. Overall, the computational results in Table 1 are remarkably consistent with experimental data on the functional status of kinase activation (based on K_{cat} values).

TABLE I SUMMARY OF COMPUTATIONALLY PREDICTED ACTIVATION SCORES AND COMPARISON WITH EXPERIMENT

Mutant	Activation Status, experiment	A-loop H Bonds Score	C-Helix H bonds Score	SASA Score ¶	Weighted Score
A1200V	WT	1/d	1/d		WT
D1349H	WT	1/d			WT
F1174L	+			1/p	A
F1245C	+			1/p	A
I1170N	+	1/p	1/p	1/p	A
I1183T	+	1/d	1/p		A
I1250T	WT				WT
L1204F	WT	1/d	1/d	1/d	WT
T1343I	WT				WT
V1229M	WT	1/d	1/d	1/d	WT
Y1278S	+	1/p	1/p	0/p	A

Column 1 describes the mutation and Column 2 summarize the activation status based on experimentally determined K_{cat} values: WT=wild type-like (< K_{cat} for wildtype), +=activating (>5-fold increase in K_{cat}). Columns 3-5 summarize the computational predictions of the activation scores. A score of 1 is given to each mutant predicted by a particular method to be activating. The scores are further qualified by p=proximal and d=distal based on how proximal/distal the given mutation is to the given sub-region of the protein (indicated by the column title). Column 6 provides the weighted score by factoring the individual scores in columns 3-5 and by proximity status.

V. CONCLUSION

We classified the type of the mutation (hydrophilic to hydrophobic, polar to non polar etc.) and its location (A-loop, C-helix, P-loop, N-loop etc.) and ascribed a mechanism-based functional significance of the mutation on the kinase activation. Based on our recent work [1], we already have demonstrated the success of this approach on the effect of mutations on the activation mechanisms of ErbB family kinases. Analyzing the effects of each activating mutation on ALK protein dynamics described here helps to reveal how the mutation functionally changes the intramolecular interactions within the kinase. The 25 mutations identified by Maris et al. are found in different regions of the TKD [3], and different activating mutations have been shown to have different sensitivities to inhibitors [2]. Our studies presented here show that not all mutations have the same mechanism of constitutive activation. Collectively, the results are helpful in the rational design of mutant-specific inhibitors and to rationalize the effect of mutation on inhibitor (Crizotinib) sensitivity in a given cohort of patients. Our results also enable the delineation of molecular events surrounding the activation mechanisms thereby complementing traditional experiments, which often do not provide the molecular context to ongoing clinical and biochemical studies. We expect that the current study on ALK with suitable validation will transform our computational approach to enable future predictions of driver oncogenic mutations with low false-positive rates, and can hence serve an important in silico tool toward personalized cancer therapy.

ACKNOWLEDGMENT

We thank Mark Lemmon, Yael Mosse, and John Marris for insightful discussions throughout this work.

REFERENCES

[1] A. J. Shih, S. E. Telesco, R. Radhakrishnan, "Analysis of somatic mutations in cancer: molecular mechanisms of activation in the ErbB family of receptor tyrosine kinases", *Cancers*, vol. 3, no. 1, pp. 1195-1231, March 2011.

[2] S. C. Bresler, A. C. Wood, E. A. Haglund, J. Courtright, L. T. Belcastro, J. S. Plegaria, K. Cole, Y. Toporovskaya, H. Zhao, E. L. Carpenter, J. G. Christensen, J. M. Maris, M. A. Lemmon, Y. P. Mossé, "Differential inhibitor sensitivity of anaplastic lymphoma kinase variants found in neuroblastoma." *Sci. Transl. Med.*, vol. 3, no. 108, p. 114, Nov 2011.

[3] J. M. Maris, Y. P. Mosse, J. P. Bradfield, C.Hou, S. Monni, R. H. Scott, S. Asgharzadeh, E. F. Attiyeh, S. J. Diskin, M. Laudenslager, C. Winter, K. A. Cole, J. T. Glessner, C. Kim, E. C. Frackelton, T. Casalunovo, A. W. Eckert, M. Capasso, E. F. Rappaport, C. McConville, W. B. London, R. C. Seeger, N. Rahman, M. Devoto, S. F. Grant, H. Li, H. Hakonarson, "Chromosome 6p22 locus associated with clinically aggressive neuroblastoma," *N. Engl. J. Med.*, vol. 358, no. 24, pp. 2585-93, Jun 2008.

[4] C. C. Lee, Y. Jia, N. Li, X Sun, K. Ng, E. Ambing, M. Y. Gao, S. Hua, C. Chen, S. Kim, P. Y. Michellys, S. A. Lesley, J. L. Harris, G. Spraggon, "Crystal structure of the ALK (anaplastic lymphoma kinase) catalytic domain," *Biochem. J.*, vol. 430, no. 3, pp. 425-37, Sep 2010.

[5] A. J. Shih, S. E. Telesco, S. H. Choi, M. A. Lemmon, R. Radhakrishnan, "Molecular dynamics analysis of conserved hydrophobic and hydrophilic bond interaction networks in ErbB family kinases," *Biochem. J*., vol. 436, no. 2, pp. 241-251, Jun 2011.

[6] S. E. Telesco, R. Radhakrishnan, "Atomistic insights into regulatory mechanisms of the HER2 tyrosine kinase domain: a molecular dynamics study," *Biophys. J*., vol. 96, no. 6, pp. 2321-2334, Mar 2009.

Hybrid Model for Tumor Spheroids with Intratumoral Oxygen Supply Heterogeneity*

Georgios Tzedakis, Eleftheria Tzamali, Vangelis Sakkalis, Alexandros Roniotis, and Kostas Marias, *Member, IEEE*

***Abstract*— Tumor growth involves numerous biochemical and biophysical processes related to the invasion of surrounding tissue and metastasis. Such phenomena occur at different scales of time and space. The desire to understand the interactions of these complex processes has given rise to various computational models allowing for multiple variable modeling using continuous, discrete and the most recent hybrid approaches. This paper presents a hybrid mathematical model of solid tumor invasion that incorporates both continuous macro-scale and discrete cell-level descriptions. Cell-based description reflects individual cell movement and state, while the continuous part formulates the nutrient supply of the tumor. In the presented model, apart from the usage of homogenous oxygen supply, intratumoral nutrient sources are introduced.**

I. Introduction

PRIMARY solid tumors are thought to arise from nodes of cells that have mutated in certain key genes. The key difference of the transformed cells is that they proliferate in an unsupervised way. As the tumor increases in size, the uncontrolled alterations lead to more complex tumor behaviors, such as metastasis and angiogenesis. The tumor at its initial stage or when a new metastasis emerges goes through simpler avascular growth with nutrient being supplied from the surrounding healthy tissue. When tumor cells exceed the capacity of the medium to support further growth, blood vessels from the pre-existing vascular network are recruited. Most models try to incorporate either the early avascular phase or the more complex vascular phase [1].

In order to perform *in silico* experiments there is a need to study the various phases and scales describing different levels of biocomplexity. Discrete event mapping allows for the description of individual cell dynamics according to some stochastic rules and are mainly applicable at the sub cellular and cellular levels, whereas continuum approaches more often assume that solid tumor behavior can be predicted in terms of its global interaction with the surrounding and underlying tissue properties and a few internal parameters related to the proliferation rate [2][3][4][5].

In this work we present a hybrid model where tumor growth is described using a discrete approach, under homogenous nutrient supplies formulated using continuous mathematics. Additional sources centered on selected outermost points of the tumor introduce introtumoral heterogeneity. This is done to show the existence of dependency between the variance in the nutrient concentration and the tumor morphology. Also such process could describe the transition from avascular to early vascular phase.

*This work was supported by the community initiative Program INTERREG III, Project "ΥΠΕΡΘΕΝ", financed by the European Commission through the European Regional Development Fund and by National Funds of Greece and Cyprus.

G. Tzedakis, E. Tzamali, V. Sakkalis, A. Roniotis and K. Marias are with the Institute of Computer Science, FORTH, Vassilika Vouton, GR-70013 Heraklion, Crete, Greece (gtzedaki; tzamali; sakkalis; roniotis; kmarias}@ics.forth.gr).

II. Background

As stated earlier, various models have been introduced in an effort to research or predict tumor growth. Some models utilize continuous equations to describe tumor growth, while others describe cancer cells as discrete automata-like entities. The key difference is that the former group of models provides qualitative results while the latter group offers quantitative results. Continuous models do not permit analysis of small-scale structures and patterns. Discrete models, being harder to rescale, are deemed inappropriate when large cell populations are concerned. Hybrid models are favored due to their ability to enlist the best elements out of both worlds [6]. Hybrid models use the continuous part to describe the tumor environment and the discrete part to describe individual cancer cells. It should be noted however that each of the aforementioned models are often developed independently from different specialized research groups even if there is a possibility of linking currently fragmented approaches increasing the potential of examining cancer under a global prism that consolidates crucial information from different levels of complexity [7].

III. Model

The model presented in this paper is classified as hybrid. It consists of a continuous part and a discrete part, which constantly interact. The continuous part describes the microenvironment around the tumor cells, while the discrete part aims to model the cancer cells. With fixed time step τ, models are consecutively updated. The continuous part is executed first driving the discrete part that is executed next. Both parts are described in the following sections.

A. Continuous Model part

In most hybrid models the continuous part is used to describe the tumor environment [8][9]. It is well known that cancer cells need nutrients and oxygen to grow and invade. It is assumed that oxygen is the only limited substrate critical for growth and proliferation. In addition, it is assumed that (avascular phase) oxygen is produced at the same rate by every healthy tissue. In other words, sources exist in places not populated by cancer cells. Oxygen diffuses and decays naturally into the extracellular matrix, but is consumed by the tumor cells with an increased rate. The resulting diffusion equation that describes oxygen supply is:

$$\frac{\partial c}{\partial t} = D_c \nabla^2 c + \beta S - \gamma N - \alpha c$$

where $c, D_c, \beta, \gamma, \alpha$ are positive constants describing the oxygen concentration, the oxygen diffusion coefficient, production rate, cancer uptake and natural decay rates, respectively. Term N can be 0 or 1, representing the absence or, respectively existence of a living tumor cell, the term S can be any number but we limit its use to integer values, by introducing a weight for the source at each individual grid point.

The equation is solved on a $[0, L] \times [0, L]$ square grid using no-flux boundary conditions. The parameter values are rendered dimensionless in such a way that a square of the grid represents 25 μm (approximately the size of a single cell). This is important for the validity of the interactions between the continuous and discrete parts of the model. Consumption rate β is determined in such a fashion that the overall oxygen flow remains stable through the tumor growth, resembling the growth of tumor spheroids. We ensure that the total oxygen production rate will be constantly $\tilde{\beta} L^2$, the rate with which oxygen is provided when only healthy tissue exists. Thus, β is determined by:

$$\beta = \frac{\tilde{\beta}\, L^2}{\sum S_{i,j}}$$

B. Discrete Model part

The discrete modeling part of the presented hybrid model assumes the following:

- A normal square lattice describes the area of interest.
- Every grid point is considered to be a cell.
- Healthy and dead cells are ignored.
- Living cancer cells can have two states, active or quiescent. The active state is the state at which the cell is preparing for proliferation. The quiescent state means that the cell has been prepared for proliferation but cannot yet proliferate due to lack of free neighbors.
- All cancer cells are arbitrarily chosen for execution of the discrete life cycle.

A flow diagram describing the different steps of the model execution is depicted in Figure 1.

The first step of the tumor cell life cycle is to check whether the lattice point at hand has enough oxygen. If the oxygen is below a viable threshold then the cell dies. Dead cancer cells, like healthy cells, are treated as empty space (cells can proliferate and replace them) but are not considered as sources of oxygen. In the case that the cell has sufficient oxygen, then its age is increased by τ hours.

If the current age of the cell passes the proliferation age threshold then the cell "attempts" to proliferate, if it fails to proliferate it enters the quiescent state. Failure to proliferate occurs if there is not enough space for the cell to proliferate to one of its eight neighbors. In the case that the proliferation successfully takes place (one or more neighboring free spaces are found) the proliferation age, in our case 16 hours, is subtracted from the cell age and an identical copy of itself is placed at a free space. The location for the new cell is chosen randomly among the neighboring free spaces.

If a cell stays in the quiescent state, due to being unable to proliferate, it does not age since it already has been prepared for proliferation. Note, however, that it will die if the oxygen drops below the viability threshold. The only way for a cell to leave the quiescent state is if a neighboring empty space is made available in order to proliferate.

If a quiescent cell finds space to proliferate then (since its age is reduced) it is considered active and prepares to proliferate again. Apart from following a slightly different life cycle from active cells, quiescent cells also consume half the oxygen that is consumed by their active counterparts.

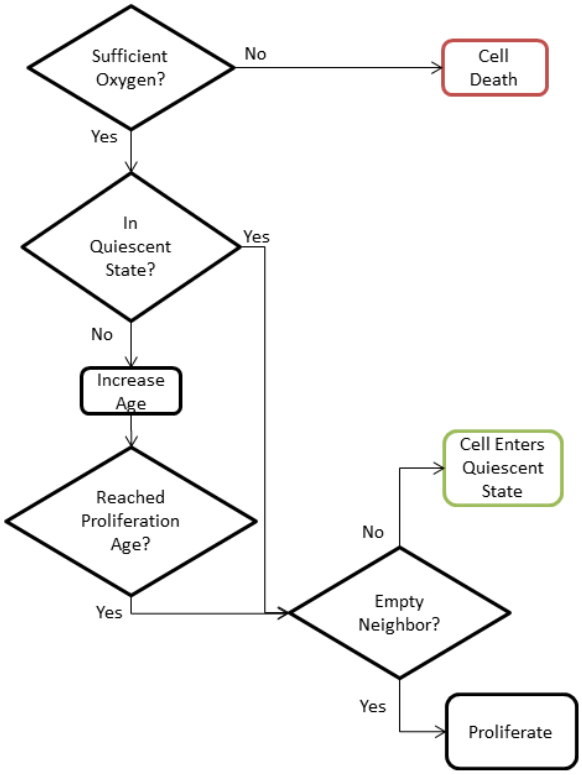


Figure 1. A flow chart that describes the life cycle steps of the discrete tumor cell

IV. Implementation

We have implemented the described model in Math Works™ Matlab (version 7.1). As there is no analytical solution to the particular diffusion equation, we approximated the solution numerically. In particular, for the solution of the equation, finite differences were used (Crank-Nicolson scheme) [10][11]. To ensure numerical accuracy and stability, the time step used was one fifth of the spatial step. To effectively approximate the solution of the sparse linear systems of equations that are derived when implementing the Crank-Nicolson, the Minimal Residual method was used.

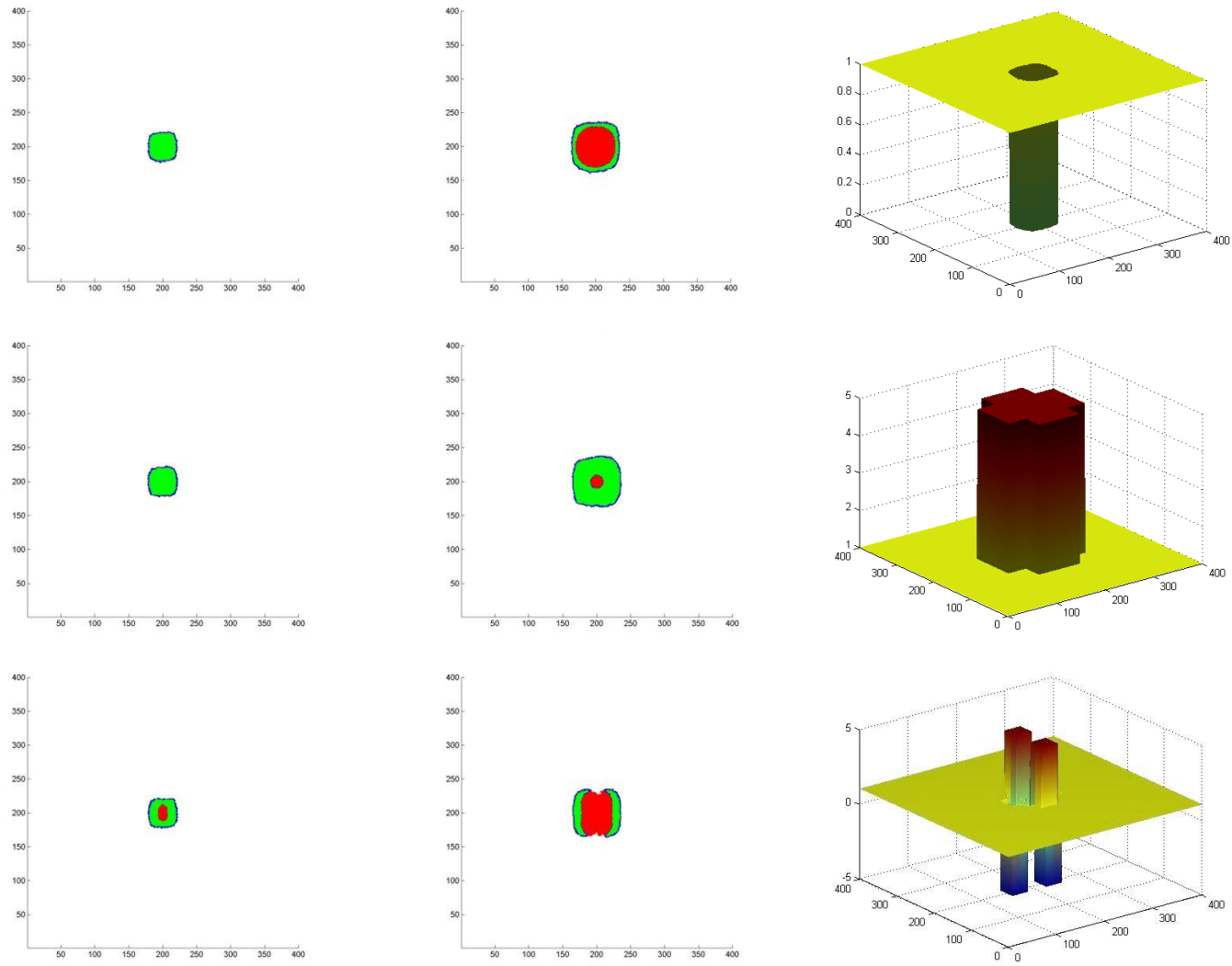


Figure 2. Each row contains the results from the respective experiment: the first row contains the results from the first experiment, the second row from the second experiment and the third row from the third experiment. The first column depicts cancer cells after 15 iterations (5 fictious days). The central column illustrates cancer cells after 30 iterations (10 fictious days). Blue, green and red coloured cells indicate active living cells, quiescent living cells and dead cells, respectively. The third column represents the source/sink weights after 30 iterations.

V. RESULTS

Our experiments aim at modeling the difference between the avascular and the early vascular phase and studying how tumor morphology is affected. In the avascular phase, the cancer receives the nutrients from the environment at a steady rate, uniformly from all the directions. In the vascular phase, due to angiogenesis the nutrient flow starts to depend on the vascularity, in other words it becomes less homogenous. The parameter values used are shown in Table I. Initial cancer population was 225 cells positioned on the center of the grid for every scenario. As a precaution, to avoid the simultaneous proliferation of all the cells, the initial population received random initial ages varying from 0 to 8 hours (the discrete time hump value).

The avascular phase is modeled by adding every healthy cell as a source of weight of 1 in the set of sources S. The vascular phase is modeled by adding sources/sinks to reduce the homogeneity in the oxygen production.

The first experiment (Fig. 2 - 1st Row) assumes that the sources are equivalently distributed on the healthy tissues. This describes the homogneous oxygen supply observed in the avascular phase. On the top left corner of Figure 2 we can see the cancer after 15 iterations (5 fictious days). The blue cells are the active proliferating ones and the green cells are the cells in the quiescent state. The center image at top row shows the cells after 30 iterations. Red coloured cells are the

TABLE I. PARAMETER VALUES

Symbol	Description	Value
L	Grid size	400
τ	Discrete time jump	8 hours
D_c	Oxygen diffusion parameter	0.25
$\tilde{\beta}$	Healthy tissue initial oxygen production rate	6.25
γ	Acive cancer cell oxygen uptake	161
α	Oxygen decay	0.0125
	Cancer cell proliferation age	16 hours
	Viable oxygen threshold	.05

cells that have died due to lack of oxygen. The top right cornered figure shows the source weights. It is evident that after 30 iterations (10 fictious days) the live cancer has formed a rim of averagely 8 cells.

As a second experiment (Fig. 2 – 2nd Row), after the tumor cell population exceeds 250 cells, 4 sources are added to create less homogenous oxygen concentrations. In an attempt to model the fact that through the process of angiogenesis new sources are created on the outer layers of the cancer, the locations of the sources move as the cancer expands on the grid. The sources are centered at the points

$$(iS1, L/2), (iS2, L/2), (L/2, jS1) \text{ and } (L/2, jS2)$$

where $iS1, iS2, jS1$ and $jS2$ are the left, right, top and bottom edges that the cancer has proceeded at that given time. The range r of the sources depends on the number of the living cancer cell population p and is given by:

$$r = 1 + \lceil p/100 \rceil$$

where $\lceil . \rceil$ denotes the ceiling function. Thus, since the sources are square shaped, each source is comprised of $2r + 1$ grid points. The weight of each extra source point is 5. In other words, the oxygen production rate of these sources is five times the rate of the normal sources while the total flow remains the same. The second row of figure 2 shows the respective values for the second trial. On the rightmost image the population depedent weighted sources are illustrated. In comparison to the first experiment, after 30 iterations the necrotic core is smaller but the tumour morphology remains the same. The average value of the living cell rim radius is increased to 24.

The third experiment (Fig. 2 – 3rd Row) assumes the same sources described previously. The difference this time is that the first two sources have weight of -5 effectively turning them into sinks. The final row of figure 2 shows the cancer population on the left and central illustrations after 15 and 30 iterations respectively. It can be observed that the tumor morphology very soon begins to vary from the previous spherical to a shape that resembles an ellipse. The respective weighted sources after 30 iterations can be seen in the bottom right corner.

VI. Conclusion

In our experiments we showed that solid tumor morphology can be affected by tampering with the nutrient source morphology when different weighted sources are applied on a normal grid. One important aspect is that the total oxygen flow rate remained unaffected.

Future work includes the addition of interaction points between the discrete and the continuous model components. Such interactions include the local nutrient values affecting individual cell proliferation time and oxygen consumption rate. Beneficial for the discrete model would also be the addition of more cell states with varying behaviors and the ability of cell movement to adjacent grid points. Finally, we plan to apply filters to the nutrient values before feeding them to the aforementioned moving procedure, in order to force the appearance intricate structures of tumor morphology, e.g. fingering patterns that appear in aggressive and metastatic real-life tumors.

References

[1] D. Hanahan, R.A. Weinberg, .*The hallmarks of cancer*. Cell, vol. 100, no. 1, pp. 57-70, 2000.

[2] A. Roniotis, V. Sakkalis, K. Marias, I. Karatzanis and M. Zervakis, *In-depth analysis and evaluation of diffusive glioma models*, IEEE Trans. Inf. Tech., Vol. 16, No. 3, pp. 299-307, 2012.

[3] V. Sakkalis et al., *Evaluation framework for the multilevel macroscopic models of solid tumor growth in the glioma case*, Conf Proc IEEE Eng Med Biol Soc. 2010, pp. 6809-12, 2010.

[4] A. Roniotis, G. Manikis, et al. *High grade glioma diffusive modeling using statistical tissue information and diffusion tensors extracted from atlases*, IEEE Trans. Inf. Tech., Vol. 16, No. 2, pp. 255-263, 2011.

[5] P.P. Delsanto et al., *A multilevel approach to cancer growth modeling* J. Theor. Biol., Vol. 250, No 1, pp. 16-24, 2007.

[6] K. A. Rejniak and A.R.A. Anderson, *Hybrid Models of Tumor Growth*. Wiley Interdisciplinary Reviews: Systems Biology and Medicine, vol. 3, no. 1, pp. 115–125, 2011.

[7] S. Sfakianakis, V. Sakkalis, K. Marias et al., *An Architecture for Integrating Cancer Model Repositories*, Conf Proc IEEE Eng Med Biol Soc. 2012, pp. 6828-31, 2012.

[8] A.R.A. Anderson, *A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion*. Math. Med. Biol 2005.

[9] P. Hinow, P. Gerlee, L. J. McCawley, V. Quaranta, M. Ciobanu, S. Wang, J. M. Graham, B. P. Ayati, J. Claridge, K. R. Swanson, M. Loveless, and A. R., Anderson, *A spatial model of tumor-host interaction: application of chemotherapy*. Math Biosci Eng, vol. 6, no. 3, pp. 521-46, 2009.

[10] A. Roniotis, K. Marias, V. Sakkalis, and G. Stamatakos. *Comparing finite elements and finite differences for developing diffusive models of glioma growth*, Conf Proc IEEE Eng Med Biol Soc. 2010, pp. 6797-6800, 2010.

[11] Roniotis, A., Marias, K., Sakkalis, V. et al., *Simulating radiotherapy effect in high grade glioma by using diffusive modeling and brain atlases*, J Biom Biot, in press 2012.

Modeling the Interplay Between Pathological Angiogenesis and Solid Tumor Growth: the Anti-angiogenic Treatment Effect*

Katerina D. Argyri, Dimitra D. Dionysiou and Georgios S. Stamatakos, *Member, IEEE*

***Abstract*— In this paper, a previous continuum approach describing vascular tumor growth under angiogenic signaling [1] is developed and extended via the inclusion of bevacizumab pharmacokinetics. The modeling approach to the problem addressed includes *inter alia* the building of the model (selection of equations, related assumptions, coupling with a pharmacokinetic model tailored to the bevacizumab paradigm, implementation and numerical solution) as well as a study of the vascular tumor growth model with results for free growth and an intermittent bevacizumab mono-therapy schedule.**

I. Introduction

Just in USA, an estimated 39920 breast cancer deaths, 160340 lung cancer deaths and 20130 lymphoma deaths are expected to occur in 2012. No further advocating is necessary for one to prove the urgent need for various scientific fields, including *in silico* oncology which is broadly considered among the cutting edge research directions of medical science, to contribute as soon as possible practical solutions to the most challenging health problem that has existed for apparently most of human history.

In order for a tumor mass to grow beyond a specific critical size (1-2 mm in diameter), it must develop a blood supply network and that is the exact goal of angiogenesis. Indeed, tumor cells secrete tumor angiogenic factors such as vascular endothelial growth factor (VEGF) in response to hypoxia and, following a cascade of biological events triggered by the aforementioned factors, the tumor is eventually penetrated by vessels. Once angiogenesis has obtained its goal, new vessels provide the tumor mass with nutrients and oxygen, which are clearly of vital importance for its survival and for extra growth. Hence, the tumor may soon reach a population of 10^9 cancer cells [2] and eventually form metastases in distant organs.

Due to the high complexity and the strongly multi-scale character of the tumor induced angiogenesis, mathematical modeling seems to be a valuable approach in qualitatively understanding the phenomenon and subsequently in supporting the treatment procedure. Even as early as the 70's the research community had already attempted to model angiogenesis both in a physiological and a pathological context mostly as a sole biological process [3][4] despite the fact that anti-angiogenic treatment had already been proposed as a theoretical therapeutic approach [5]. However, with the appearance in the 90's of actual experimental data suggesting that angiogenesis blockage could lead to tumor regression [6], the first completed mathematical models accounting for the interaction of tumor growth and tumor induced angiogenesis had already been introduced by the end of the decade [7][8]. During the first years of the new century and up to this point in time much progress has been done toward developing mathematical models which describe malignant tumor growth and take into account tumor induced angiogenesis explicitly [9-11] or even implicitly as perturbator of the proliferation related parameters that characterize the tumor (e.g. through the modification of the probability of a newborn cell to re-enter the cell cycle in relation to its value in the largely necrotic layer of the tumor) [12][13].

In the aforementioned context, the structure of the paper is the following: In section II, a modeling approach describing vascular tumor growth under angiogenic signaling is outlined and briefly analyzed. In paragraph A the biological processes incorporated into the model as well as the underlying assumptions of the specific approach are stated, in paragraph B the modeling approach by Poleszczuk and his coworkers [1] that was used as a basis for our approach is introduced and in paragraph C the modification and extension of [1] by coupling it with another model simulating bevacizumab pharmacokinetics [14] is described. In section III, the implementation of the two coupled models is outlined and in section IV indicative results and respective comments are presented. Finally, the paper concludes with section V, where the main points of our work as well as suggestions and extensions are discussed.

II. Vascular Tumor Growth under Anti-angiogenic Treatment: A Continuum Approach

The model that was selected as a starting point to our approach was the one introduced by Poleszczuk *et al*. [1], a member of the wider family of approaches that were based on the scientific work of Hahnfeldt *et al*. [7].

A. Biological Mechanisms Addressed by the Model and Underlying Assumptions

The continuum approach that is presented in paragraph B accounts for tumor cell proliferation, tumor cell apoptosis, postvascular dormancy (the state where angiogenesis stimulation and inhibition has come into balance), endothelial cell death, spontaneous loss of functional vasculature, excretion of endogenous proangiogenic factors such as vascular endothelial growth factor, fibroblast growth factors, etc., excretion of endogenous antiangiogenic factors e.g.

* This work has been supported in part by the European Commission under the projects TUMOR: Transatlantic Tumor Model Repositories (FP7-ICT-2009.5.4-247754) and p-Medicine: Personalized Medicine (FP7-ICT-2009.5.3-270089).

K.D. Argyri (kargyri@mail.ntua.gr), D.D. Dionysiou (dimdio@esd.ece.ntua.gr) and Georgios S. Stamatakos (corresponding author) (phone: +30 210772 2287; fax: +30 2107733557; e-mail: gestam@central.ntua.gr) are with the Institute of Communication and Computer Systems, National Technical University of Athens, 9 Iroon Polytechniou, GR 157 80, Zografos, Greece

.

angiostatin, endostatin, etc., antiangiogenic treatment induced endothelial cell death and resulting tumor cell death.

The underlying assumptions made in the context of the basic framework of the model are that the tumor is a three dimensional spheroid, that the diffusion process is in a quasi-stationary state (i.e. that the tumor growth rate as well as the rate of change of drug concentration are relatively small compared to the rate of distribution of angiogenesis stimulators) and that the concentration of the stimulator/inhibitor is a radially symmetric function i.e.,

$$\|x\| = \|y\| \Rightarrow n(x) = n(y) \quad (1)$$

where $n(\cdot)$ is the function that gives the concentration of the stimulator/inhibitor.

B. Description of the Dynamical System in [1]

The model makes use of a variable carrying capacity i.e. the maximum tumor volume that can be supported by the given vasculature, a concept originally introduced in [7]. The dynamical system described in [1] is consisted of a pair of ordinary differential equations which reflect the interplay between tumor volume (V) and carrying capacity (K).

Equation (2) describes the rate of change of the tumor volume V. As it is also the case in [7], the authors have adopted Gompertzian growth equation:

$$\frac{dV}{dt} = -\lambda_1 \cdot V \cdot \ln\left(\frac{V}{K}\right) \quad (2)$$

Equation (3) describes the rate of change of the carrying capacity K:

$$\frac{dK}{dt} = -\lambda_2 \cdot K + c \cdot \frac{\left(\beta + V^p\right)}{\alpha \cdot \left(\beta + V^p\right) + I(t)} - d \cdot K \cdot V^{2/3} \quad (3)$$

The parameters and variables involved in (2) and (3) with their units are explained in table I.

TABLE I. PARAMETERS AND VARIABLES INVOLVED IN (2),(3)

	Mathematical Symbol	Units	Description
Independent variable	t	time	time
Dependent Variable	V	vol	tumor volume
Dependent Variable	K	vol	tumor capacity
Parameter	λ_1	1/time	proportionality constant
Parameter	λ_2	1/time	proportionality constant
Parameter	c	1/(time · vol^p) · conc	proportionality constant
Parameter	d	1/(time · $vol^{2/3}$)	proportionality constant
Parameter	α	conc /vol^p	constant
Parameter	β	vol^p	constant
Parameter	p	≥0	Hill coefficient
Parameter	I	conc	drug concentration in plasma

In order to compute the exact forms of the 2nd and 3rd terms of (3) the authors applied a diffusion-consumption equation for the concentration of stimulators as well as for the concentration of inhibitors.

C. Incorporation of Anti-angiogenic Treatment Pharmacokinetics: the Paradigm of Bevacizumab Mono-therapy

This paragraph describes the computation of antiangiogenic drug concentration $I(t)$ in a given time-point t, a function that is involved in equation (3). We will proceed with the case of Bevacizumab which probably is the most popular representative of the wider family of antiangiogenic agents in clinical practice.

The specific agent is a recombinant humanized monoclonal antibody that binds vascular endothelial growth factor (VEGF), the main mediator of tumor angiogenesis, thereby inhibiting the interaction of VEGF with its receptors (Flt-1 and KDR) on the surface of endothelial cells.

Relevant literature strongly suggests that two – compartmental models assuming first-order elimination give the best description of bevacizumab pharmacokinetic data [15]. For the needs of reliable simulation of anti-angiogenic treatment we have studied and incorporated into [1] the two-compartmental pharmacokinetic model proposed by Metzler and coworkers [14].

If T stands for the duration of infusion, the function that describes bevacizumab concentration in plasma is the following:

While the infusion takes place, namely while $0 \le t < T$:

$$I(t) = \frac{k_A}{k_e \cdot V_c} \cdot \left(1 - R \cdot e^{-at} + Q \cdot e^{-bt}\right) \quad (4)$$

After the infusion, namely while $T \le t$:

$$I(t) = \frac{k_A}{k_e \cdot V_c}\left(Q \cdot \left(e^{-b \cdot t} - 1\right) \cdot e^{-b(t-T)} - R \cdot \left(e^{-a \cdot T} - 1\right) \cdot e^{-a(t-T)}\right) \quad (5)$$

The parameters and variables involved in (4) and (5) with their units are explained in table II.

TABLE II. PARAMETERS AND VARIABLES INVOLVED IN (4), (5)

	Mathematical Symbol	Units	Description
Independent variable	t	time	time
Dependent variable	I	conc	concentration of administered inhibitor in blood
Parameter	n	-	number of infusions
Parameter	T	time	duration of infusion
Parameter	V_c	vol	volume of central compartment
Parameter	k_{12}	1/ time	inter-compartmental rate constant (from central compartment to

			outer compartment)
Parameter	k_{21}	1/time	inter-compartmental rate constant (from outer compartment to central compartment
Parameter	k_e	1/time	elimination rate constant
Parameter (Internal)	*dosage*	conc	dosage
Parameter	k_A	mass/time	rate of infusion
Parameter (Internal)	*w*	mass	patient's weight
Parameter	t_i , $i = 1,\ldots,n$	time	administration time-points of anti-angiogenic treatment

* *a*, *b*, *Q*, *R* are computed based on inter-compartmental rate constant values

III. Implementation - Numerical Solution

We proceeded with the implementation of the model in MATLAB. The m-files involved in the implementation of the model are listed and explained below.

DrugConcentration.m: This function implements the two-compartmental pharmacokinetic model [14] for administration via intravenous infusion (as it is the case for bevacizumab in clinical practice). It calculates the concentration of the antiangiogenic agent at each time point taking into account remnants of previous infusions, if any.

VascularTumorGrowth.m: This function computes the derivatives involved in (2), (3) which describe the vascular tumor growth.

VascularTumorGrowth_main.m: Given initial values (V_0, K_0) and a time interval, the specific script file resolves the problem with the solver ode45, which implements a Runge Kutta method with a variable time step for efficient computation. It also plots the variables *V* and *K* (representing tumor volume and tumor capacity respectively) as functions of time in the same system of axes as well as the phase – plane of the system.

It is noted that the experiments of free growth and intermittent treatment were performed on a desktop computer with an AMD Phenom(tm) II X6 1055T Processor 2.80 GHz, 8.00GB RAM and with the 64 – bit version of Windows 7 and that the code execution time is of the order of a few seconds.

IV. Indicative Results

Multiple code executions have been conducted testing the code with different parameter values, time span, initial values and code module (free growth, constant treatment, intermittent treatment). Below, two indicative results are presented by adopting the classical graph representation.

A. Free Growth

In this execution, the results of which are exhibited in Fig. 1, λ_1 was set as equal to 0.003465 1/day in order to reflect a typical value of human breast cancer doubling time i.e. 200 days and the initial values as equal to (V_0, K_0) = (200 mm^3, 4000 mm^3). A plateau of the tumor volume is obtained approximately after 3000 days, the value of which is determined by the point where endogenous angiogenesis stimulators and inhibitors come into balance. The tumor is characterized by the parameter values that are listed in table III.

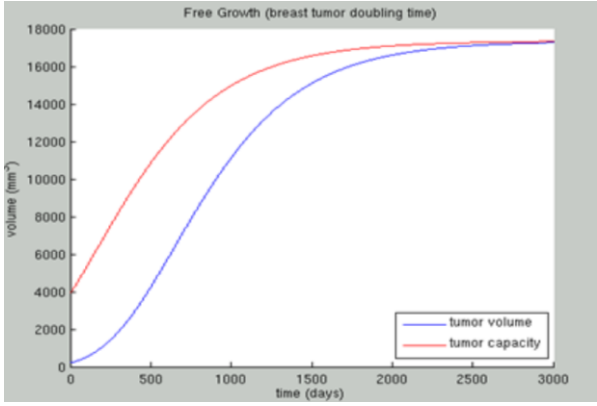


Figure 1. Simulation results for an untreated breast tumor characterized by the parameter values shown in table III

B. Intermittent Treatment

Fig. 2 presents the time course of volume *V* and carrying capacity *K* for a tumor treated with bevacizumab according to a real-life treatment scheme applied in clinical practice [15]. The tumor and the administered treatment are characterized by the parameter values shown in table III. Treatment begins the 300th day when the tumor size is already equal to 8000 mm^3 i.e. when the radius of the tumor is equal to 20 mm. Before the 300th day an initial overshoot of tumor volume (*V*) is observed due to the fact that the tumor is untreated. With the start of the regimen administration a sharp decrease of tumor capacity is obvious and then the response of tumor volume follows. This makes sense as anti-angiogenic treatment influences directly the vascular compartment and through it, the tumor compartment responses. Another interesting observation is that each time the curve *K*(*t*) intersects curve *V*(*t*), the monotonicity of the function *V*(t) changes. This observation is reasonable if one keeps in mind that tumor capacity is actually defined as the maximal tumor volume sustainable by the current vasculature.

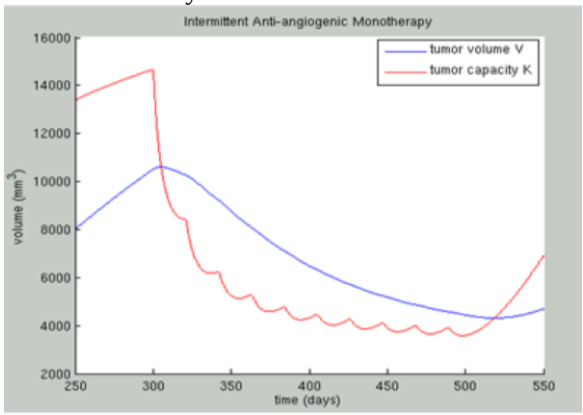


Figure 2. Simulation results for a tumor treated with bevacizumab according to a scheme used in clinical practice [15] and characterized by the parameter values shown in table III

TABLE III. PARAMETER VALUES USED IN EXHIBITED CODE EXECUTIONS

Code Executions			
	A	***B***	*Units*
n	n/a	9	-
T	n/a	1/16	days
t_i , $i = 1,\ldots,n$	n/a	t_1=300, t_2=321, … t_9=468	days
Vc	n/a	2390	ml
k_{12}	n/a	0.223	1/day
k_{21}	n/a	0.215	1/day
k_e	n/a	0.0779	1/day
dosage	n/a	15	mg/ml
k_a	n/a	calculated internally	mg/day
w	n/a	50	kg
λ_1	0.003465	0.013	1/day
λ_2	0	0	1/day
c	5.85	5.85	mg/(ml·day · mm^{3p})
d	0.00873	0.00873	1/(day · mm^2)
α	1	1	mg /(mm^{3p} · ml)
β	1	15	mm^{3p}
p	0	0	-

V. CONCLUSION

The main objective of this paper is to briefly present a mathematical model not only describing the interplay between tumor induced angiogenesis and solid tumor growth but also simulating the therapeutic effect of a real-life intermittent bevacizumab mono-therapy schedule. The model is currently being clinically and experimentally adapted and validated. A thorough sensitivity analysis is in progress in order to address the uncertainty of the model parameters. Multiple code executions have already been conducted testing the code with various parameter values, simulation time span, initial values and code modules. A meaningful way to extend this work would be to proceed to modifications so as to come up with a model simulating combination treatment consisted of both cytotoxic and anti-angiogenic agents.

ACKNOWLEDGMENT

The authors would like to convey thanks to E.Ch.Georgiadi and E.A. Kolokotroni for their invaluable assistance and support.

REFERENCES

[1] J. Poleszczuk, M. Bodnar, U. Foryś, "New approach to modelling of antiangiogenic treatment on the basis of Hahnfeldt et al. model", *Math Biosci Eng.*, vol. 8, no. 2, pp. 591-603, April. 2011.

[2] N. V. Mantzaris, S. Webb and H. G. Othmer, "Mathematical modeling of tumor-induced Angiogenesis", *J. Math. Biol*, vol. 49, pp. 111–187, 2004.

[3] D.F. Zawicki, R.K. Jain, G.W. Schmid-Schoenbein and S. Chien, "Dynamics of neovascularization in normal tissue", *Microvasc. Res.*, vol. 21, pp. 27–47, 1981.

[4] D. Balding, and D. L. S. McElwain, "A mathematical model of tumour-induced capillary growth", *J. theor. Biol*, vol. 114, pp. 53–73, 1985.

[5] J. Folkman, "Tumour angiogenesis: therapeutic implications". *N. Engl. J. Med.*, vol. 285, pp. 1182–1186, 1971.

[6] M. S. O'Reilly, L. Holmgren, C. Chen and J. Folkman, "Angiostatin induces and sustains dormancy of human primary tumours in mice". *Nat. Med.*, vol. 2, pp. 689–692, 1996.

[7] P. Hahnfeldt, D. Panigrahy, J. Folkman and L. Hlatky, "Tumour development under angiogenic signaling: A dynamical theory of tumour growth, treatment response and postvascular dormancy", *Cancer Res.*, vol. 59, pp. 4770-4775, 1999.

[8] E. De Angelis and L. Preziosi. "Advection–diffusion models for solid tumour evolution in vivo and related free boundary problem", *Math. Models Methods Appl. Sci.*, vol. 10, pp. 379–408, 2000.

[9] F. Billy, B. Ribba, O. Saut, H. Morre-Trouilhet, T. Colin, D. Bresch, J.P. Boissel, E. Grenier, and J.P. Flandrois., "A pharmacologically based multiscale mathematical model of angiogenesis and its use in investigating the efficacy of a new cancer treatment strategy", *Journal of Theoretical Biology*, vol. 260, no.4, pp. 545-562, 2009.

[10] A. d'Onofrio and A. Gandolfi, "A family of models of angiogenesis and antiangiogenic anticancer therapy", *Mathematical Medicine and Biology*, vol. 26, pp. 63-95, 2009.

[11] M. Kohandel, M. Kardar, S. Sivaloganathan and M. Milosevic, "Dynamics of tumor growth and combination of anti-angiogenic and cytotoxic therapies", *Phys. Med. Biol.*, vol. 52, pp. 3665, 2007.

[12] G.S. Stamatakos, D.D. Dionysiou, E.I. Zacharaki, N.A. Mouravliansky, K.S.Nikita, N.K. and Uzunoglu , "In silico radiation oncology: combining novel simulation algorithms with current visualization techniques" , I*EEE Proceedings: Special Issue on Bioinformatics: Advances and Chalenges* , vol. 90, no. 11 , pp. 1764-1777 , 2002.

[13] D.D.Dionysiou, G.S. Stamatakos, N.K.Uzunoglu, K.S.Nikita and A. Marioli , "A Four Dimensional In Vivo Model of Tumour Response to Radiotherapy: Parametric Validation Considering Radiosensitivity, Genetic Profile and Fractionation", *J. theor. Biol.*, vol. 230, pp. 1-20, 2004.

[14] C. M. Metzler,"Usefulness of the two-compartment open model in pharmacokinetics", *J. Am. Stat. Assoc*., vol. 66, pp. 49-53, 1971.

[15] J-F Lu, R. Bruno, S. Eppler, W. Novotny, B. Lum and J. Gaudreault, "Clinical pharmacokinetics of bevacizumab in patients with solid tumours", *Cancer Chemother Pharmacol*, vol. 62, pp. 779–786, 2008.

Prostate tumor growth modeled by a statistically –modulated PUN scheme

Caterina Guiot, Elisabetta Garibaldi, Domenico Gabriele, and Pietro Gabriele

Abstract— **Prostate tumor (PT) normally has an indolent course during the first 10 to 15 years. Tumor progression and its clinical management, in the elderly patients, is therefore becoming a social problem, and the feasibility of non-therapeutic approaches, such as the Wait & Watch (W&W) one, requires a better understanding of the tumor natural history. Our model aims at connecting physical, biological and statistical knowledge in order to validate empirical views and to propose to clinicians a simple and effective 'therapy-simulator', based on a large experimental sample, helping them in the difficult goal of personalizing patients' therapy.**

I. Introduction

Prostate tumor (PT) normally develops slowly during the first 10 to 15 years and is often diagnosed late in life. Its clinical management, particularly in the elderly patients, is progressively becoming a social problem, and the choice of the best therapeutic approach, included the Wait & Watch (W&W) , requires a better understanding of the tumor natural history [1, 2].

Modeling approaches to prostate tumor growth proposed in the last decades were predominantly either rigorous, mathematically based and relying on meaningful biological parameters (but often not very realistic…) ones or, more recently, epidemiological and statistical models, which take profit from the huge amount of data now available for their validation but are often more empirical and not solidly based on physical and biological reasoning. A remarkable example of interconnection between the two views is given by Dimonte's paper [3], (see also its comprehensive bibliography for the above references).

One of the key points characterizing a model, and its ability to be applied in clinical settings, is the choice of its parameters. To make validation easier, prostate tumor progression can be parametrized by means of the 'TNM' clinical staging, as an indirect evaluation of the actual tumor volume.

Growth rate and tissue carrying-capacity will be related to other two sensible and clinically relevant parameters, i.e. the Prostate-Specific Antigen (PSA) and the Gleason score (GS).

C.G. is with the Dept. of Neuroscience, University of Torino, Torino, 10125 Italy, (phone: +39-11-6708166; fax: +39-11-6708174; e-mail: caterina.guiot@unito.it).
E. G. and P.G. are with *IRCC* - Institute for Cancer Research and Treatment, D. O. Radiotherapy, 10060 Candiolo, Torino
(e-mail: elisabetta.garibaldi@ircc.it, pietro.gabriele@ircc.it)
D. G. is graduate student at University of Torino (e-mail: gabrielepda@alice.it)

PSA plays a crucial role in the early detection of the PT [4] and in the monitoring of therapeutic failures [5]. However, it is sometimes a confounding parameter because of its poor specificity, being PSA currently produced by the prostate and increased also in many non-cancerous diseases like prostate hyperplasia or inflammation and even recent mechanical stimulation. We therefore, differently from Dimonte's paper [3], will account for PSA (and GS) not as a given 'input' parameter but as a factor which modulates the possible scenarios of PT evolution.
We propose an approach which assumes a 'non-stocastic' growth law, which is very flexible and able to incorporate the tumor growth 'spurts' related to the well known invasive and metastatic development by accordingly 'remodulating' the values of the parameters. Such remodulation is performed by combining a 'deterministic' model of the natural history of the tumor with a statistical description, which mirrors the actual individual variability in tumor growth, and can be validated by clinical data.

In what follows we will describe: a) the mathematical model and the main parameters which characterize its description of tumor growth**;** b) some typical 'scenarios' for prostate tumor progression; c) the implementation of the therapies and the validation strategies. The aim is that of providing clinicians with a sort of 'simulator' helping optimization of the therapeutic approach with an 'intelligent' use of previous clinical experiences.

II. Outlook of the Model

We describe tumor growth using the recently proposed [6] Phenomenological Universalities (PUN) approach, already applied to a wide range of topics [7]. In order to describe the PUN methodology from an applicative point of view, let us start with the first order nonlinear growth differential equation

$$\frac{dy}{dt} = a(y)y, \tag{1}$$

where $y(t)$ represents the variable of interest (e.g. the tumor mass or dimension or the number of cells) and $a(y)$ the (unknown) growth rate. To integrate Eq.(1), it is necessary to make some assumption about the rate a or, equivalently, on its time derivative or acceleration b , e.g. assuming that b is given by an expansion in a , i.e.

$$\beta = \sum_{i=1}^{N} c_i \alpha^i = \beta\alpha + \gamma\alpha^2 + \ldots. \qquad (2)$$

We call UN the class generated by the solution of the coupled ordinary differential equations (ODE) (1) and (2), when in the latter only the first N terms are considered. The functions $y(t)$ that one obtains for the first UN classes ($N = 0,1,2$) have a very wide range of applications. In fact:

A. for $N = 0$, i.e. U0, $\beta = 0$; $y(t)$ represents a self-catalytic or exponential growth function. By integrating over the two ODEs, we obtain

$$y(t) = y_0 \exp(\alpha_0 t), \qquad (3)$$

where $\alpha_0 = \alpha(t = 0)$ is the initial growth rate Here and in the following we normalize the variable $y(t)$, so that $y(0) = 1$. This function describes an unlimited growth, and is obviously too simple and unrealistic for tumors.

B. for $N = 1$, i.e. U1, $\beta = \beta\alpha$ and

$$y(t) = y_0 \exp\left[\alpha_0/\beta\left(\exp(\beta t) - 1\right)\right]. \qquad (4)$$

Equation (4) represents the Gompertz law [8], which has been extensively used in all kinds of growth problems for almost two centuries, and especially to tumors [9]. The parameter β represents a factor which is related to the so called carrying capacity of the system , which may be defined as $\beta = -\alpha_0 / \ln y_\infty$, where y_∞ is the asymptotic value of $y(t)$.

C. for $N = 2$, i.e. U2, $\beta = \beta\alpha + \gamma\alpha^2$ and

$$y(t) = y_0 \exp\left[1 + \alpha_0\gamma/\beta\left(1 - \exp(\beta t)\right)\right]^{-1/\gamma}, \qquad (5)$$

which yields a generalization of West's law [10].

A more accurate description of the tumor growth process should, however, be taken into account for periods of discontinued growth, such as dormancies or sudden increases in growth rate, which are normally experienced. This can be achieved by superimposing on an overall U1 or U2 model one or more "mathematical spurts", modeled as small additional corrections that are also to be described in the framework of a UN class.

Provided that the overall growth curve can be fitted, up to a certain level of approximation, by Eq. (4), to restrict our treatment to the class U1 and limit the number of "free" parameters of the model only to β, since y_0 and α_0 may be directly evaluated from the first few data.

We then add on Eq. (4) a number M (usually one or two) of "spurts", each given by

$$\Delta y_m(t) = G(t_m, \sigma_m, t) y_{0m} \exp\left\{\left(\frac{\alpha_m}{\beta_m}\right)\left[\exp\left(\beta_m (t - t_m)\right) - 1\right]\right\} \qquad (6)$$

Being y_{0m} i the amplitude of the spurt at $t = t_m$, and $G(t_m, \sigma_m, t)$ i the Gauss cumulative function with average value t_m and standard deviation σ_m, defined by

$$G(t_m, \sigma_m, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{(\tau - t_m)^2}{2\sigma_m^2}} d\tau \qquad (7)$$

A special case of the Gauss cumulative function is the Heaviside function

$$H(t_m, t) = G(t_m, 0, t) = \begin{cases} 0 & \text{if } t \le t_m; \\ 1 & \text{if } t > t_m. \end{cases} \qquad (8)$$

The corresponding PUN classes will be called ($U1+GM$) and ($U1+HM$), respectively, with $M = 1,2,\ldots$ being the number of spurts assumed.

III. Scenarios for prostate cancer

Anatomically 5 distinct lobes are distinguishable in prostate gland: Anterior (A) Medium (M), Posterior (P) that is reperable at rectal touch, Right Lateral (RL) and Left Lateral (LL). [11] We can actually envisage a Progression Scheme, where the prostate tumor grows initially within a glandular lobe surrounded by intact capsule. According to Mc Neal [12] 4 zones are of interest: the transitional or periuretral zone where benign prostatic hypertrofia and 20% of tumors normally develops, the peripheral zone, where about 70% of tumors origin, the central zone (25%) and the anterior zone (10%). Tumors confined within a lobe are normally scored T1 or T2 (see Table 1).

TABLE I: DETAILS OF THE T CODES FROM TNM CLINICAL CLASSIFICATION FOR PROSTATE CANCER

	TNM Clinical Classification for PROSTATE CANCER
TX.	Primary tumor **cannot be assessed**
T0.	**No evidence** of primary tumor
T1	Clinically inapparent tumor **not palpable or visible by imaging**
T1a	Tumor incidental histological **finding in 5% or less of tissue resected**
T1b	Tumor incidental histological **finding in more than 5% of tissue resected**
T1c	Tumor **identified by needle biopsy** (e.g., because of elevated PSA)
T2	Tumor confined within prostate
T2a	Tumor involves **one half of one lobe or less**
T2b	Tumor involves **more than half of one lobe**, but not both lobes
T2c	Tumor involves **both lobes**
T3	Tumor **extends through the prostatic capsule**
T3a	**Extracapsular extension** (unilateral or bilateral)
T3b	Tumor **invades seminal vesicle(s)**
T4	**Tumor is fixed or invades adjacent structures** other than seminal vesicles: bladder neck, external sphincter, rectum, levator muscles, or pelvic wall

T3 score defines the invasion of the full organ and T4 the loco-regional invasion. Such events would obviously occur with a different timing depending, for instance, on their different geometrical positions inside the prostate . Provided it is centrally positioned, a longer time will be required to reach the boundary of the organ and to start invading the surrounding. The statistical distribution of the times is therefore described by the 'average time' tmin Eq. (6,7) for organ filling, while the two 'tails', described by σm in Eq (6,7) account for the cases in which the tumor seed is just near to (or very far from) the boundary. Due to the rich vascular net of the prostate (both arterial, prostatic a. and intern pudenda, and venous (uretral plexus, vescical-prostatic bilateral pudend plexus, hypogastric veins [11] as well as the extended lymphatic drainage, tumor seeds can easily colonize the regional lymphatic nodes (score N1) or the distant nodes (score M1a), the bones (M1b) or other distant organs (M1c). The actual timing of near and distant invasions are again different for individuals, depending on factors which can be taken into account only statistically. Furthermore, the likelihood of spreading beyond confines of the prostate is strictly related to both GS and PSA : for instance according to Mayo Clinic Prostate Cancer Guide [13] 12% of low-grade tumors (GS: 2-4) spread beyond prostate in 10 years, 33% of medium-grade tumors (GS:5-6) spread beyond prostate in10 years and 61% of high-grade tumors (GS: 7-10) spread beyond prostate in 10 years. Also PSA dosage, which is considered normal if < 2,5 ng/mL before 50 yy and < 4 ng/mL in elderly patients, was shown to be lower then 10 ng/mL in confined TP, then 20 ng/mL in lymphonodal involvements and to be >20 ng/mL in metastatical spread [14].

Note that during the follow-up after therapy the normality limits are 0,2 ng/mL after radical prostatectomy and Nadir + 2 ng/mL after radiotherapy.

Actually, some combination of pretreatment prognostic factors, such as clinical stage, PSA and GS, has been empirically related to the extraprostatic spread of PT with nomograms such as the Roach and Memorial Formulas (widely used by Radiation Oncologists) or the Partin Table (widely used by surgeons) [15-19].

Different scenarios should be therefore envisaged for different initial values of GS and PSA, by differentiating at least 3 x 4 = 12 different scenarios.

As an example, a possible scenario pertaining to the case of low GS and low PSA will be described by Fig. 1 , where the average times and the Gaussian amplitudes are to be defined according to the available data.

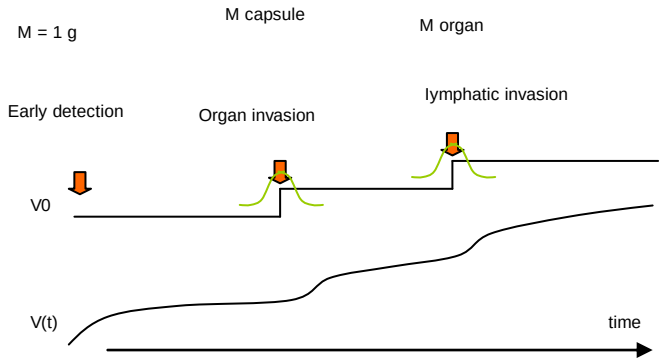


Figure 1: **sketch of the expected tumor growth progression in a low GS and PSA scenario.**

IV. IMPLEMENTATION OF THERAPIES

Tumor detection is normally followed by the choice of some single or multi-therapeutic strategy.

In recent years, due to the fact that most T1-T2 cancers have an indolent course during the first 10 to 15 years and patients are often middle-aged, the so-called Watchful Waiting (W&W) approach is preferred.

When stadiation and/or GS and PSA values suggest a more incisive intervention, various approaches are normally selected between surgery (radical prostatectomy classic or robotic with/without lymphadenectomy), radiotherapy (brachytherapy / external radiotherapy with 3D/IMRT/IGRT), hormons , chemotherapy or other therapies (criotherapy, HIFU etc.) only for recurrent disease

All the above therapies can be described by proper mathematical algorithms which accounts for timing, dosage, one or multi-shot (see for instance [20, 21]).
An extension of the model presented in part I will be provided based on the data actually available.

V. Validation strategies

Data will be provided by three main ways: the first are the tumor epidemiological registers available in Turin city and Biella city of Regione Piemonte , which report diagnosis, some epidemiological data of patients, treatments and the outcome ; the second are some clinical registers of Radiotherapy centers as the IRCC in Candiolo (TO), the Mauriziano Hospital in Turin and the Radiotherapy center in Biella (BI) and the urology center in S. Giovanni Bosco Hospital in Turin ; lastly the final way are data coming from National Institute Tumori in Milan where a W&W protocol for early prostate cancer is currently applied.

The subgroups of patients who didn't undergo any therapy or followed the W&W strategy will be used to validate the natural history of the prostate cancer. The other patients will be divided according to the mono (surgery or radiotherapy) or multi-therapies (surgery + radiotherapy or hormons + radiotherapy) and their data will be used to validate the corresponding part of the model.

VI. Conclusion

First of all, the natural history of the PT will be investigated by comparing model scenarios and clinical data. Particular attention will be devoted to various 'empirical' formulas, such as the Roach, Partin and Memorial H. nomograms [16, 18, 19], to be validated on the basis of the model. In a second stage the impact of therapies on the PT progression will be evaluated.
The ideal goal of such a model would be that of providing clinicians of an affordable but simple tool for producing some 'simulation' of the clinical expectations for any single patients making an 'intelligent' use of the previous available experiences. After model validation, a further prospective study in which the model best predictions will be compared with current clinical managed decisions, possible performed on a blind, multicenter and randomized recruiting base, is envisaged.

.

References

[1] Johansson JE, "Natural History of Early, Localized Prostate Cancer", *JAMA* 2004 June, pp 2713-2719

[2] Lyratzolpoulus G, "Population based time trends and socioeconomic variation in use of radiotherapy and radical surgery for prostate cancer in a UK region: continuous survey", *BMJ* 2010 n 340, p 1928.

[3]G. Dimonte, "A cell kinetics model for prostate cancer and its application to clinical data and individual patients", *J. Theoret. Biol* May 2010, pp 420-442 November 2009, pp 8507-8515.

[4] Welch HG, "Prostate cancer diagnosis and treatment after the introduction of prostate-specific antigen screening: 1986–2005", *J Nat Cancer Inst* 2009 n 101, pp 1325–9.

[5] Roach M 3rd, "Defining biochemical failure following radiotherapy with or without hormonal therapy in men with clinically localized prostate cancer: recommendations of the RTOG-ASTRO Phoenix Consensus Conference", *Int J Radiat Oncol Biol Phys* 2006 Jul, pp 965-74.

[6] Castorina P ,Guiot C , "Classification scheme for phenomenological universalities in growth problems in physics and other sciences.". *Phys Rev Lett* , 2006 May vol 96, pp 188-201

[7] Delsanto, PP; Gliozzi, AS.; Guiot, C. "Scaling, growth and cyclicity in bioogy: a new computational approach". *Theor Biol & Med Model* , 2009 Feb vol 5, 5 DOI: 10.1186/1742-4682-5-5

[8] B. Gompertz, "On the nature of the function expressive of the law of human mortality and on a new mode of determining life contingencies" . *Phil. Trans. Roy. Soc. Lond.*, 1825 vol 123, p 513 .

[9] Norton L, "The growth curve of an experimental solid tumor following radiotherapy", *J. Natl. Cancer Inst*. 1977 June, Vol 58, pp 1735-1741.

[10] West GB, "A general model for ontogenetic growth", *Nature* 2001 Oct, vol 413, pp 628-631.

[11] Atlas of Clinical Anatomy, F. Netter

[12] McNeal. The Zonal Anatomy of the Prostate, *JE – Prostate* 1981 n 2, pp 35-49

[13] Mayo Clinic.com. Prostate Cancer Guide. Available at: http://www.mayoclinic.com/health/prostate-cancer/PC99999.
Prostate Cancer in California. Ed. Mill PK. Public Health Institute. 2000.

[14] Proust-Lima C, "Determinants of change in prostate-specific antigen over time and its association with recurrence after external beam radiation therapy for prostate cancer in five large cohorts", *Int J Radiat Oncol Biol Phys* 2008 n 72, pp 782–791.

[15] D'Amico AV, "Pretreatment nomogram for prostate-specific antigen recurrence after radical prostatectomy or external-beam radiation therapy for clinically localized prostate cancer", *J Clin Oncol* 1999 Jan, pp 168-72.

[16] Johns Hopkins Medicine. The Partin Tables. Available at http://urology.jhu.edu/prostate/partintables.php. Accessed October 19, 2010

[17] http://urology.ihu.edu/prostate/partintables.php, "Combination of prostate-specific antigen, clinical stage and Gleason score to predict pathological stage of localized prostate cancer. A multi-institutional up-date". *JAMA* 1997 n 277, pp 1445-1451.

[18] Roach M III, "Predicting the risk of lymph node involvement using the pre-treatment prostate specific antigen and Gleason score in men with clinically localized prostate cancer", *Int J Radiat Oncol Biol Phys* 1994 n 28, pp 33-37

[19] Memorial Sloan-Kettering Cancer Center. Prostate Cancer Nomograms. Available at: http://www.mskcc.org/cancer-care/adult/prostate/prediction-tools. Accessed October 19, 2010

[20] Castorina P, C. Guiot, " Tumor growth instability and its implications for chemotherapy", *Cancer Res.* November 2009, pp 8507-8515.

[21] Castorina P., Gabriele P., Guiot C, "Growth laws in cancer: implications for radiotherapy", *Radiation Res* Sept 2007, vol 168, pp 349-356

The Continuous Mathematics Based Glioblastoma Oncosimulator: Application of an Explicit Three Dimensional Numerical Treatment of the Skull-Glioblastoma Neumann Boundary Condition on Real Anatomical Data*

Stavroula G. Giatili and Georgios S. Stamatakos, *Member IEEE*

***Abstract*—The Continuous Mathematics Based Glioblastoma Oncosimulator is a platform for simulating, investigating, better understanding, and exploring the natural phenomenon of glioma tumor growth. Modelling of the diffusive-invasive behaviour of glioma tumour growth may have considerable therapeutic implications. A crucial component of the corresponding computational problem is the numerical treatment of the adiabatic Neumann boundary conditions imposed by the skull on the diffusive growth of gliomas and in particular glioblastoma multiforme (GBM). In order to become clinically acceptable such a numerical handling should ensure that no potentially life-threatening glioma cells disappear artificially due to oversimplifying assumptions applied to the simulated region boundaries. However, no explicit numerical treatment of the 3D boundary conditions under consideration has appeared in the literature to the best of the authors' knowledge. Therefore, this paper aims at providing an outline of a novel, explicit and thorough numerical solution to this problem. Additionally, a brief exposition of the numerical solution process for a homogeneous approximation of glioma diffusion-invasion using the Crank – Nicolson technique in conjunction with the Conjugate Gradient system solver is outlined. The entire mathematical and numerical treatment is also in principle applicable to mathematically similar physical, chemical and biological diffusion based spatiotemporal phenomena which take place in other domains for example embryonic growth and general tissue growth and tissue differentiation. A comparison of the numerical solution for the special case of pure diffusion in the absence of boundary conditions with its analytical counterpart has been made. In *silico* experimentation with various adiabatic boundary geometries and non zero net tumour growth rate support the validity of the corresponding mathematical treatment. Through numerical experimentation on a set of real brain imaging data, a simulated tumour has shown to satisfy the expected macroscopic behaviour of glioblastoma multiforme, on concrete published clinical imaging data, including the adiabatic behaviour of the skull. The paper concludes with a number of remarks pertaining to the potential and the limitations of the diffusion-reaction approach to modelling multiscale malignant tumour dynamics.**

* This work has been supported in part by the European Commission under the projects ContraCancrum: (FP7-ICT-2007-2- 223979), TUMOR: Transatlantic Tumor Model Repositories (FP7-ICT-2009.5.4-247754) and p-Medicine: Personalized Medicine (FP7-ICT-2009.5.3-270089).

S. G. Giatili is with *In Silico* Oncology Group, Laboratory of Microwave and Fibre Optics, Institute of Communication and Computer Systems, National Technical University of Athens. 9, Iroon Polytechniou, Zografos, GR-157 80, Greece (e-mail: giatili@mail.ntua.gr).

G. S. Stamatakos is with *In Silico* Oncology Group, Laboratory of Microwave and Fibre Optics, Institute of Communication and Computer Systems, National Technical University of Athens. 9, Iroon Polytechniou, Zografos, GR-157 80, Greece (corresponding author phone: + (30) 210 772 2287; fax:+(30)2107723557; e-mail: gestam@central.ntua.gr).

I. Introduction

The Oncosimulator is a concept of integrative cancer biology, a complex algorithmic construct, a biomedical engineering system, and eventually in the future a clinical tool that primarily aims at supporting the clinician in the process of optimizing cancer treatment in the patient – individualized context through conducting experiments *in silico* i.e. on the computer. Additionally, it is a platform for simulating, investigating, better understanding, and exploring the natural phenomenon of cancer, supporting the design and interpretation of clinicogenomic trials and finally training doctors, researchers, and interested patients alike [1].

Glioblastoma multiforme (GBM) is a very aggressive glioma and a classical example of a highly invasive and diffusive tumour. GBM cell diffusion in the brain is a reasonable first approximation of the migration of glioma cells along structures such as the basement membranes of blood vessels or the glial limitans externa that contain extracellular matrix (ECM) proteins. Frequentlty, invasive glioma cells are also found to migrate along myelinated fiber tracts of white matter. Despite improvements in cancer treatment, its overall prognosis is still very poor. Due to its markedly diffusive character, a significant component of the tumour cannot be delineated based on standard tomographic imaging techniques such as CT, MRI and PET. Glioma cells are found in tissues surrounding the tumor, even after total resection of the tumor parts detectable by MRI scanning.This constitutes an important limitation to the optimal design of both surgical excision and therapeutic irradiation of the tumour. In order to partly alleviate the problem, mathematical modelling of diffusive tumour growth has been proposed. To this end a number of diffusion-reaction based models dealing primarily with the morphology of tumour growth have been developed [2 – 4].

In accordance with the diffusion-reaction based approach, the tumour is considered a spatiotemporal distribution of continuous cell density which follows the general diffusion-reaction law. Tumour growth can be expressed by the following statement [2,3]:

Rate of change of tumour cell population= diffusion (motility) of tumour cells + net proliferation of tumour cells - loss of tumour cells due to treatment

The macroscopic formulation of diffusion, leads to a partial parabolic differential equation. A single tumour cell may constitute the initial tumour within a three-dimensional medium. In the case of glioma, the simulated region of interest may include part of the skull which acts as an adiabatic boundary for the diffusion of the brain tumour, precluding migration beyond it. Subsequently, the simulation of tumor growth and invasion can be viewed as a boundary value problem strongly dependent on the values assigned on the physical boundary of the definition domain. As a result, the mathematical treatment of the biophysical processes taking place in the vicinity of anatomic boundaries must satisfy specific constraints. Zero flux boundary conditions have to be applied on the anatomic boundaries of the skull surface. Thus if Ω is the brain domain on which the diffusion equation is to be solved the previous statement can be symbolically formulated through the following differential equation [2]:

$$\left\{\begin{array}{l} \dfrac{\partial c}{\partial t} = \nabla\cdot(D\nabla c) + \rho c - G(t)c \quad in\ \Omega \\ c(\vec{x},0) = f(\vec{x}), \quad initial\ condition \\ \hat{n}\cdot D\nabla c(\vec{x},t) = 0 \ \ on\ \ \partial\Omega,\ Neumann\ Boundary\ Condition \end{array}\right\} \quad (1)$$

The variable c denotes the cell concentration at any spatial point defined by the position vector $\vec{x}$ and time t. The parameter D denotes the diffusion coefficient and represents the active motility of tumour cells. The term ρ represents the net rate of tumour growth including proliferation, loss and death, $\hat{n}$ is the unit vector normal to the boundary $\partial\Omega$ of the domain Ω and $f(\vec{x})$ is a known function that defines the initial spatial distribution of malignant cells. The term $G(t)$ accounts for the temporal profile of treatment and as a first facilitating approximation $G(t)=k$ may be assumed constant. The latter may crudely model a continuous administration of radiation e.g. through special radioisotope based implants. A more realistic assumption is to assign $G(t)$ different values for different time intervals reflecting various chemotherapeutic and/or radiotherapeutic schedules. The simulation domain R of which Ω is a subdomain is defined as:

$$R=\{(x, y, z)|\ a<x<b,\ s<y<d,\ e<z<f\}. \quad (2)$$

II. Explicit Numerical Formulation Of The Boundary Conditions

The first step in the problem approximation process is to make spatial and time discretization introducing a computational grid that is applied on the anatomic region of interest. Having defined the computational grid, the diffusion component of the problem is numerically solved using the Crank Nicolson numerical method in conjunction with the Conjugate Gradient method. The Crank-Nicolson method is second-order accurate in both time and space and unconditionally stable. The non - stationary iterative Conjugate Gradient method (CG) is widely used for the solution of linear sparse systems. In a spatiotemporal model of tumor growth, any anatomic boundaries of the space where tumor can grow are very important for the outcome of the simulation. In order to complete the model formulation for clinical glioblastoma growth homogeneous Neumann boundary conditions have to be added.

The glioma invasion problem involves an irregularly shaped domain. Therefore, a biologically meaningful solution has to allow for the investigation of a wide range of elementary local domain geometries. Several specific cases have been examined in order to address the geometry of the irregularly shaped skull boundary. For each boundary mesh node (lying at the center of the multi-colored structure of Fig.1 (and therefore not visible) all its 6 adjacent nodes (lying towards all the $x+$, $x-$, $y+$, $y-$, $z+$, $z-$ directions) are considered in order to numerically apply the boundary condition on it i.e. on (x_i, y_j, z_k) [5].

The boundary condition according to (1) is:

$$\hat{n}\cdot D\nabla c = 0 \quad on \quad \partial\Omega\,. \quad (3)$$

In order to evaluate the boundary condition for each grid point (x_i, y_j, z_k) and maintain the block tridiagonal structure of the coefficient matrix $\vec{A}$ of the resulting linear system of algebraic equations $\vec{A}\vec{x} = \vec{b}$ and the second-order accuracy of the approximation we introduce a "fictitious node" into the computational grid. The "fictitious node" produces an extra row of unknowns in the computational grid. Evaluating the boundary condition at each boundary grid point (x_i, y_j, z_k) yields six equations mathematically similar (but not identical) to the equation corresponding to the following case:

At the boundary grid point (x_i, y_j, z_k) in the negative x direction $x-$:

$$-\left.\frac{\partial c}{\partial x}\right|_{(x_i,y_j,z_k)} = 0 \Rightarrow c_{i+1,j,k} = c_{F_{i-1,j,k}} \quad (4)$$

where $F_{i,j,k}$ denotes a fictitious node.

The total number of the different cases of nodes having boundary node(s) as their neighbour(s) that have been considered is 26. This has led to the formulation of 26 algebraic equations mathematically similar (but not identical) to (5). An appropriate equation out of the set of these 26 equations is used for any index triplet (i, j, k) belonging to the boundary. By fixing indices i, j, k to specific values, the 26 equations can produce all elementary boundary arrangements encountered in the case of an arbitrarily shaped boundary.

An indicative case and equation is the following:

At the boundary grid point (x_i, y_j, z_k) where the skull lies only in the positive x direction:

$$[1+6\lambda-\frac{\Delta t}{2}(\rho-G)]c_{i,j,k}^{t+1}-\lambda(2c_{i-1,j,k}^{t+1}+c_{i,j+1,k}^{t+1}+c_{i,j-1,k}^{t+1}+c_{i,j,k+1}^{t+1}+c_{i,j,k-1}^{t+1})= \quad (5)$$
$$[1-6\lambda+\frac{\Delta t}{2}(\rho-G)]c_{i,j,k}^{t}+\lambda(2c_{i-1,j,k}^{t}+c_{i,j+1,k}^{t}+c_{i,j-1,k}^{t}+c_{i,j,k+1}^{t}+c_{i,j,k-1}^{t})$$

where $\lambda=D\Delta t$ / $[2(h)^2]$ for $h=\Delta x=\Delta y=\Delta z$, $t_n=n\Delta t$, and $n=0,1,2\ldots$

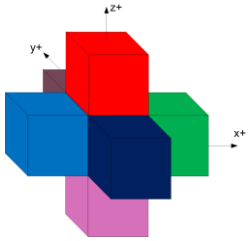


Figure 1. A boundary mesh node (lying at the center of the multi-colored structure and therefore not visible) with its 6 adjacent nodes.

III. Numerical Experiments

In order to test the numerical schemes implemented for solving (1) and support the correctness of the overall mathematical treatment presented, a number of pertinent computational scenarios have been executed. They have included inter alia: numerical checks regarding convergence and stability of the algorithm and the code, checks regarding mass conservation and linearity for the theoretical case of pure diffusion, comparison of the model with the analytical solution to the special case of an initial Gaussian cell concentration profile, spatial symmetry studies for simple symmetric geometries and numerical validation of the adiabatic behaviour of the boundary implementation.

IV. Clinical Validation Aspects

For the clinical validation of the model, a real human head has been considered. The three dimensional image depicted in panel A of Fig. 2 has been constructed using a freely available T1 weighted MRI head dataset (J. Orchard, http://www.cs.uwaterloo.ca/~jorchard/mri/).

Several snapshots of a growing virtual glioblastoma tumour corresponding to various time points are depicted in Fig. 3. It is noted that although the internal anatomy of brain is visible in the panels of this figure, homogeneous diffusion of tumour cells has been assumed within the skull cavity as a first approximation. The concentration of tumour cells within the initial simulated tumour has been arbitrarily assumed uniform and equal to 10^6 cells per cubic millimeter. The following parameter values have been used: diffusion coefficient D=0.0065 cm^2/d, [2], h=0.2cm, Δt=0.5d , and net tumour growth rate ρ=0.012 units per day [2]. It should be noted that the gross spatial pattern of glioblastoma growth, especially in the vicinity of the skull boundary, is in very good agreement with actual published clinical observations [6-7]. The change of the total number of tumor cells inside the mesh is depicted in Fig. 4.

The diameter of a sphere with volume equal to a glioblastoma tumour of fatal imageable dimensions is about 6cm [2]. The latter corresponds to a volume V_{fatal} of 113.04cm^3. In order for the tumour to increase in imageable volume from $V_{fatal}/2$ to V_{fatal}, 26 days are needed according to
the simulations. This approximation to doubling time is in good agreement with the clinically reported glioblastoma

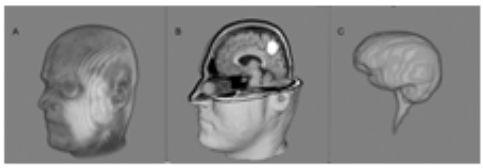


Figure 2. (A) A three dimensional rendering of a set of T1 weighted MRI slices of a real human head. (B) An initial fictitious virtual spherical glioblastoma tumor of radius equal to 1.4 cm (denoted by white color) lying inside the skull cavity. (C) Reconstruction of the brain cavity of the skull following segmentation of each MRI slice.

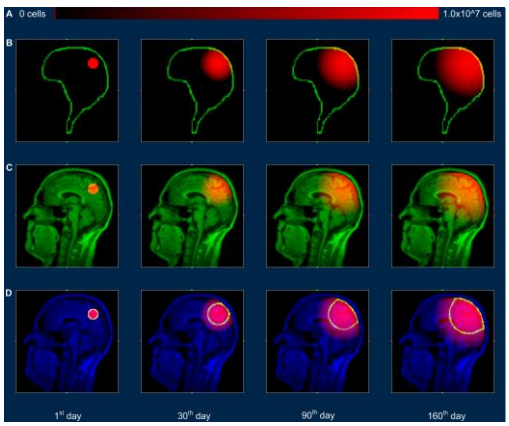


Figure 3. Schematic representation of the growth of a virtual glioblastoma tumor in vivo in sagittal planes at various time points (panel columns from left to right correspond to days 1, 30, 90 and 160 respectively). (A) The red color intensity level I depends on cell concentration according to the function $I=klog_{10}c$, where c denotes tumor cell concentration, the constant $k=255/log_{10}c_{max}$, c_{max} is the maximum value of tumor cell concentration over the entire space and time range considered during all simulations that have been included in this figure. Maximum and zero cell concentration corresponds to RGB(255,0,0) and RGB(0,0,0) respectively. (B), (C) As time increases, the tumor diffuses theoretically over the interior space of the skull cavity of the human head. (D) The yellow/bright contour defines the boundary of the tumor component in 2D that is tomographically detectable and has a cell density higher than the assumed threshold of 8000 cells per cubic millimeter.

doubling time in [8]. It is noted that in order to avoid a somewhat artificial diffusion behaviour during the first

simulated days which would be dictated by the deliberately assumed abrupt boundaries of the initial tumour, the first 14 simulated days have not been taken into account in the theoretical estimation of doubling time.

Fig. 5 depicts the relative percentage error of the total number of tumor cells calculated by comparing the predicted values generated by the model with the ones obtained by the application of the exponential growth. The exponential growth is the simplest proliferation law. The first reported work on an exponentially growing population was performed by Reverend T.R. Malthus in 1798.

A typical execution instance of the code for 6 simulated months, Δt=0.5d and for descritized mesh 130x130x130, on a 32-bit Windows Vista Platform, 4 GB RAM and processor Intel® Core™2 Duo CPU P8600 @ 2.4GHz, takes 214 sec. Further acceleration of the code execution could be achieved by using high performance computing resources.

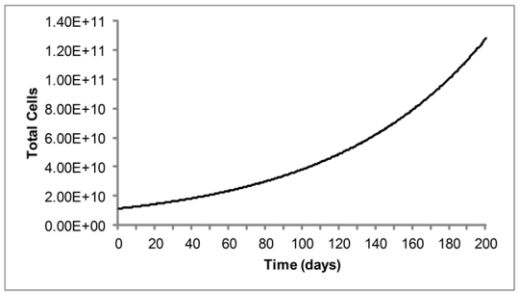


Figure 4. Total number of tumor cells for the virtual tumor

V. Conclusion

An explicit numerical treatment of the boundary conditions to be used in conjunction with a homogeneous diffusion-reaction based glioma growth model has been presented.

Systematic checking of the corresponding computer code based on numerical simulations for various adiabatic boundary geometries, zero and nonzero net tumour growth rate and zero and nonzero loss rate due to treatment have supported the validity of the corresponding mathematical treatment. Numerical experimentation on a set of real brain imaging data has demonstrated that a simulated glioblastoma tumour can satisfy the expected macroscopic behaviour of its real counterpart, including the adiabatic behaviour of the skull. The detailed treatment of the boundary conditions presented could considerably contribute to the accuracy of the solution to the diffusion-reaction equation in particular for glioblastoma tumours having their main bulk close to the skull. The composite model proposed appears to have the potential to correctly predict clinically meaningful and measurable quantities of critical importance related to the course of the disease, such as the imaging based doubling time. Obviously a strict clinical adaptation and validation procedure is a sine qua non requirement before clinical translation is envisaged. Additionally, translation or extension of the analysis presented to mathematically similar physical systems is a possibility. Extension to both inhomogeneous and anisotropic glioma diffusion is pretty straightforward. It should be noted, however, that although the continuous - finitized approach partly delineated in this paper appears to be a good choice for pure tumour growth-invasion-diffusion modelling, it seems not to be ideal for the integration of the massive multiscale biological complexity that is necessary in order to study in depth tumour response to treatment.

Discrete entity – discrete event based approaches [9-12] on the contrary have demonstrated considerable integrative potential in the context of cancer response to treatment due to the discrete character of many biological entities and features involved in this domain (e.g. discrete cell categories based on their mitotic potential such as stem cells, progenitor cells, differentiated cells; discrete cell cycle phases generally characterized by differing treatment sensitivities; discrete character of cell state transitions etc.). Nevertheless, both continuous/finitized and discrete approaches are important for the development [11] and the clinical translation [13] of the emerging oncosimulators [11].

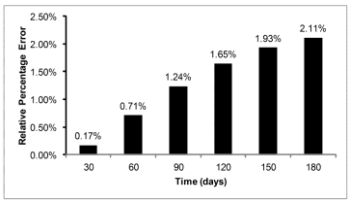


Figure 5. Plot of the relative percentage error of the total number of tumor cells for the initial tumor of Fig. 2.

Acknowledgment

Fruitful discussions with N. Graf, University Hospital of Saarland, Germany. and E. Georgiadi ICCS, National Technical University of Athens, Greece are duly acknowledged.

References

[1] G.Stamatakos "In Silico Oncology Part I: Clinically Oriented Cancer Multilevel Modeling Based on Discrete Event Simulation" In T.Deisboeck and G. Stamatakos Eds pp. 407-436 2011-01-01 CRC Press, Print ISBN: 978-1-4398-1440-6 eBook ISBN: 978-1-4398-1442-0 DOI: 10.1201/b10407-19 Boca Raton, Florida, USA, 2010/2011

[2] J.D. Murray,Mathematical Biology II, 3rd Ed., Springer-Verlag Inc.,New York, ,2003, pp. 536-613.

[3] K. R. Swanson, Quantifying glioma cell growth and invasion in vitro, Math. and Comp. Modelling. 47 (2008) 638–648.

[4] O. Clatz, P.Y Bondiau, H. Delingette, M Sermesant, S.K. Warfield, G. Malandain, N. Ayache, Brain Tumor Growth Simulation. Institute National de Recherché en Informatique et en Automatique, Rapport de recherche. 5187 (2004).

[5] S.G. Giatili and G.S. Stamatakos, A detailed numerical treatment of the boundary conditions imposed by the skull on a diffusion–reaction model of glioma tumor growth. Clinical validation aspects, Appl. Math. Comput. (2012) doi:10.1016/j.amc.2012.02.036 in press

[6] A.D. Waldman, A. Jackson, S. J. Price, C.A. Clark, T.C. Booth, D.P. Auer, P.S. Tofts, D.J. Collins, M.O. Leach, J.H. Rees, Quantitative imaging biomarkers in neuro-oncology, Nature Reviews Clinical Oncology. 6 (2009) 445-454.

[7] Glioblastoma multiforme. http://en.wikipedia.org/wiki/Glioblastoma_multiforme (2011). Accessed 26 November 2011.

[8] N.G. Burnet, R. Jena, S. J. Jefferies, S. P. Stenning, N.F. Kirkby, Mathematical Modelling of Survival of Glioblastoma Patients Suggests a Role for Radiotherapy Dose Escalation and Predicts Poorer Outcome After Delay to Start Treatment, Clinical Oncology 19 (2006) 93-103.

[9] G.S. Stamatakos ,D.D. Dionysiou, E.I. Zacharaki, N.A. Mouravliansky, K.S. Nikita, N.K. Uzunoglu, In silico radiation oncology: combining novel simulation algorithms with current visualization techniques, Proc IEEE, Special Issue on Bioinformatics: Advances and Chalenges. 90 (11) (2002) 1764-1777.

[10] D.D. Dionysiou, G.S. Stamatakos, N.K. Uzunoglu, K.S. Nikita, A. Marioli, A Four Dimensional In Vivo Model of Tumor Response to Radiotherapy: Parametric Validation Considering Radiosensitivity, Genetic Profile and Fractionation , J. Theor. Biol. 230 (2004) 1-20.

[11] G. Stamatakos, In Silico Oncology: PART I Clinically oriented cancer multilevel modeling based on discrete event simulation in: T. S. Deisboeck and G. S. Stamatakos (Eds.) Multiscale Cancer Modeling. CRC Press 2010/2011 pp 407–436, Print ISBN: 978-1-4398-1440-6, eBook ISBN: 978-1-4398-1442-0, DOI: 10.1201/b10407-19.

[12] G.S.Stamatakos, E.Ch.Georgiadi, N.Graf, E.A.Kolokotroni, D.D.Dionysiou, Exploiting Clinical Trial Data Drastically Narrows the Window of Possible Solutions to the Problem of Clinical Adaptation of a Multiscale Cancer Model, PLoS One. 2011;6(3):e17594.

[13] N. Graf, In Silico Oncology Part II: Clinical Requirements Regarding In Silico Oncology, in T. Deisboeck and G. S . Stamatakos (Eds), Multiscale Cancer Modeling, CRC Press 2010/2011, pp. 437–446, Print ISBN: 978-1-4398-1440-6, eBook ISBN: 978-1-4398-1442-0, DOI: 10.1201/b10407-2

Solving the PIHNA model while accounting for radiotherapy*

Alexandros Roniotis, Vangelis Sakkalis, Eleftheria Tzamali, Georgios Tzedakis, Michalis Zervakis, and Kostas Marias, *Member IEEE*

***Abstract*— Glioblastoma is the most aggressive type of glioma. During the last decades, several models have been proposed for simulating the growth procedure of glioma. One of the latest proposed models builds upon the proliferation – diffusion model by incorporating the angiogenic net rates and different concentration of cell populations (normoxic, hypoxic and necrotic). This proliferation- invasion- hypoxia- necrosis- angiogenesis model (PIHNA) does not take into account radiotherapy. This work presents the mathematical foundation for solving PIHNA model in two dimensions with incorporated radiotherapy effect using the Linear Quadratic Model, which uses radiobiology parameters.**

I. Introduction

GLIOMA, especially glioblastoma, is the most fatal brain cancer, despite the major advances in medicine [1-3]. The highly invasive and neoplastic growth of glioma has emerged the necessity to describe the mechanism of glioma growth, using mathematical models.

The major macroscopic models either use the diffusion reaction equation for simulating the change of tumor cell concentration (diffusive models) [4-9] or cellular automata for simulating invasion and proliferation by using deterministic cell state change rules [10-11].

This paper deals with an extended type of diffusive models for three types of cell populations and simulates radiotherapy by adding a consuming term from the Linear Quadratic Model in the respective equations.

II. Background

Diffusive models [4] simulate the change of glioma concentration in time and in space, by using two main terms for invasion (diffusion term) and proliferation (reaction term) of glioma cells. The equation used in diffusive models is the following:

$$\frac{\partial c}{\partial t} = \nabla \cdot (D\nabla c) + \rho c\left(1 - \frac{c}{k}\right) \tag{1}$$

where $c(\mathbf{x},t)$ is the is the tumor concentration in position $\mathbf{x}$ at time t, D is the diffusion coefficient (can be spatially varying), ∇ and $\nabla\cdot$ are the gradient and divergence operators respectively, ρ is the net cell proliferation rate and k is the maximum tumor cell concentration. This model assumes that cancerous cells proliferate at a constant rate which is independent of nutrients availability and the cell population has the same type.

In order to incorporate nutrient availability and oxygenation, the model of (1) can be altered to the proliferation- invasion- hypoxia- necrosis- angiogenesis (PIHNA) model [12-13]. This model includes three different types of cell population, normoxic, hypoxic and necrotic, while vasculature and angiogenic factors (e.g. VEGF) have also been incorporated into the model.

The emerging system of equations for these three types of cells, endothelial cells (vasculature) and angiogenic factors are:

$$\begin{aligned}
\frac{\partial c}{\partial t} &= \nabla \cdot (D(1-T)\nabla c) + \rho c(1-T) + \gamma hV - \beta c(1-V) - a_n nc \\
\frac{\partial h}{\partial t} &= \nabla \cdot (D(1-T)\nabla h) - \gamma hV + \beta c(1-V) - (a_h(1-V) + a_n n)h \\
\frac{\partial n}{\partial t} &= a_h h(1-V) + a_n n(c + h + v) \\
\frac{\partial v}{\partial t} &= \nabla \cdot (D_v(1-T)\nabla v) + \mu \frac{a}{K_m + a} v(1-T) - a_n nv \\
\frac{\partial a}{\partial t} &= \nabla \cdot (D_a \nabla a) + \delta_c c + \delta_h h - q\mu \frac{\alpha}{K_m + \alpha} v(1-T) - \omega av - \lambda a
\end{aligned} \tag{2}$$

where $V = v/(c+h+v)$, $T = (c+h+v+n)/k$. In first equation $c(\boldsymbol{x},t)$ is the concentration of normoxic cells at position $\mathbf{x}$ at time t. The cells diffuse at a rate D, proliferate at a rate ρ, turn to hypoxic at a rate β or turn directly to necrotic (due to contact death) at a rate a_n.

Similarly, $h(\boldsymbol{x},t)$ is the concentration of hypoxic cells, which diffuse at a rate D, turn back to normoxic at a rate γ or turn to necrotic at a rate a_h.

Continuing, $n(\boldsymbol{x},t)$ is the concentration of necrotic cells, while $v(\boldsymbol{x},t)$ and $a(\boldsymbol{x},t)$ are the vasculature (endothelial cells linked to oxygenation) and concentration of angiogenic factors, respectively. Vasculature disperses at a rate D_v, increases at a rate $a_v = \mu \frac{a}{K_m + a}$ and turns to necrotic cells at rate a_n.

Lastly, angiogenic factors are produced by normoxic cells at rate δ_c and by hypoxic cells at rate δ_h, decay at a rate λ and are washed out by vessels at rate ω. In the last two equations, μ is the proliferation rate of endothelial cells, K_m is the Michaelis – Menten constant of response of endothelial cells to angiogenic factors and q is the consumption of angiogenic factors per endothelial cell proliferation.

* This work was supported in part by the community initiative Program INTERREG III, Project "ΥΠΕΡΘΕΝ", financed by the European Commission through the European Regional Development Fund and by National Funds of Greece and Cyprus and the EC project TUMOR (FP7-ICT-2009.5.4-247754)

A. Roniotis, V. Sakkalis, E. Tzamali, G. Tzedakisand K. Marias are with the Institute of Computer Science, FORTH, Vassilika Vouton, GR-70013 Heraklion, Crete, Greece ({roniotis; sakkalis; tzamali; gtzedaki; kmarias}@ics.forth.gr). Greece (Phone: +30 2810 391672 Fax: +30 2810 391428)A. Roniotis and M. Zervakis are with the Dept of Electronic & Computer Engineering, Technical University of Crete, 73100, Chania, Greece (michalis@display.tuc.gr).

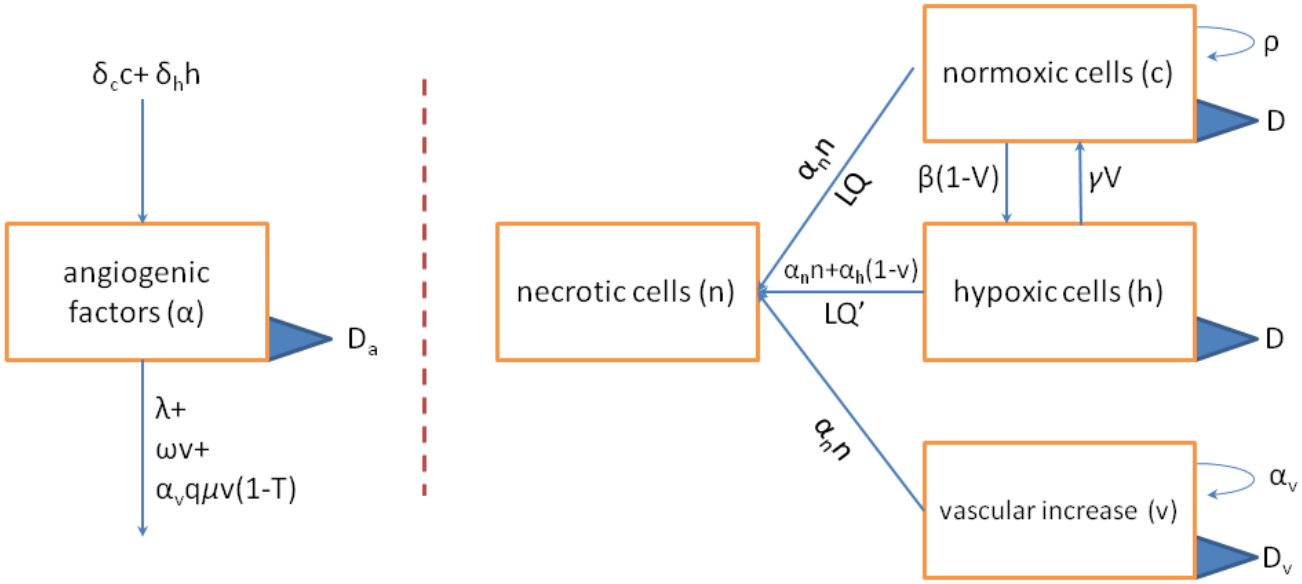


Figure 1. Modified Figure from [13]: Outline of the PIHNA-LQ mathematical interactions for the PIHNA model, by adding the radiotherapy terms extracted from the LQ model [16]. The interactions interfere normoxic cells, hypoxic cells, necrotic cells, endothelial cells (vasculature) and angiogenic factors (e.g. VEGF).

In order to incorporate radiotherapy into the PIHNA model, the Linear Quadratic Model (LQ) could be used [14-17]. According to LQ model, the probability of cells surviving S following a single dose of radiation $R(\mathbf{x}, t)$ was observed to follow the relationship:

$$S(R) = exp(-AR - BR^2) \tag{3}$$

where *linear* A and *quadratic* B are the radiobiology parameters, which are interpreted biologically as repairable single and lethal double-strand breaks to the cell's DNA, respectively [15]. In general, cancers with high motility, like glioblastoma, show a high tissue response $A/B \approx 10$ [16]. Especially for photon radiation therapy, which is most commonly applied, the hypoxic cells are as much as 2 to 3 times more resistant to radiation damage than normoxic [18]. Thus, Oxygen Enhancement Ratio (OER) has to be included in the model [19].

III. Methods

This paper investigates the incorporation of radiotherapy effect in the PIHNA model. Radiotherapy could be interpreted as an additional term for turning normoxic cells and hypoxic cells to necrotic cells in the first two equations of (2). Given the radiotherapy dose, the effectiveness of irradiation on hypoxic cells, compared to normoxic cells, is lower and this is expressed through OER. Thus, if we assume that all normoxic cells are affected by irradiation and OER is the Oxygen Enhancement Ratio, then the first three equations of (2) can be altered to the PIHNA-LQ model as following:

$$\frac{\partial c}{\partial t} = \nabla \cdot (D(1-T)\nabla c) + \rho c(1-T) + \gamma hV - \beta c(1-V) - a_n nc - LQc$$

$$\frac{\partial h}{\partial t} = \nabla \cdot (D(1-T)\nabla h) - \gamma hV + \beta c(1-V) - (a_h(1-V) + a_n n)h - LQ'h$$

$$\frac{\partial n}{\partial t} = a_h h(1-V) + a_n n(c+h+v) + LQ'h + LQc \tag{4}$$

where:

$$LQ = 1 - e^{-AR(t) - BR^2(t)}$$
$$LQ' = 1 - e^{-\frac{A}{OER}R(t) - \frac{B}{OER^2}R^2(t)} \tag{5}$$

Fig.1. illustrates the interactions between the elements of the PIHNA-LQ model for simulating proliferation, invasion, hypoxia, necrosis, angiogenesis and radiotherapy. The interactions have been extracted from equations (2), (4), (5). The diagram has been designed on the diagram of Figure 1 of [13], by adding the radiotherapy terms of the LQ model.

IV. Model Implementation

We have implemented the PIHNA-LQ model in MathWorks™ Matlab (version 7.1). As there is no direct algebraic solution to the equation system of (4), we estimated the solution numerically. More specifically, we approximated the spatiotemporal solution of the system of the five partial differential equations by developing numerical schemes of Finite Differences in two spatial dimensions.

In order to avoid instability issues, the implicit numerical method of Crank Nikolson was used for building the Finite Differences for a variable in an equation where this variable was subject to temporal change.

On the other hand, in order to simplify the procedure of solving, an explicit method (Forward Euler) was used for approximating the rest variables for the model. This means

TABLE I – THE PARAMETERS FOR PIHNA-LQ MODEL SIMULATION AND THEIR NONDIMENSIONALIZED VERSIONS

Parameter	Value	Nondimensionalized	Parameter	Value	Nondimensionalized
ρ	$\frac{ln2}{\tau}\left(\frac{1}{day}\right)$	$\tau\rho = ln2$	β	$\frac{\rho}{10}\left(\frac{1}{day}\right)$	$\frac{ln2}{10}$
γ	$0.05\left(\frac{1}{day}\right)$	$\tau\gamma = 0.0166$	D	$10^{-9}\left(\frac{cm^2}{s}\right)$	$\frac{\tau}{L^2}D = 3\ 10^{-5}$
a_n	$\frac{ln2}{50}\left(\frac{1}{day}\right)$	$\frac{ln2}{150}$	D_v	$10^{-9}\left(\frac{cm^2}{s}\right)$	$\frac{\tau}{L^2}D_v = 3\ 10^{-5}$
a_h	$\frac{\rho}{20}\left(\frac{1}{day}\right)$	$\frac{ln2}{20}$	D_a	$2.9\ 10^{-7}\left(\frac{cm^2}{s}\right)$	$\frac{\tau}{L^2}D_a = 8\ 10^{-3}$
μ	$\frac{ln2}{15}\left(\frac{1}{day}\right)$	$\tau\mu = \frac{ln2}{45}$	K_m	$5.75\ 10^{-7}\left(\frac{mmol}{cm^3}\right)$	$\frac{K_m}{a_{max}} = 10^{-3}$
δ_c	$7.59\ 10^{-16}\left(\frac{mmol}{cell\ day}\right)$	$\frac{\tau k}{a_{max}}\delta_c = 0.44\ 10^{-4}$	λ	$\frac{ln2}{64}\left(\frac{1}{min}\right)$	$\tau\lambda = 7.5\ ln2$
δ_h	$1.43\ 10^{-12}\left(\frac{mmol}{cell\ day}\right)$	$\frac{\tau k}{a_{max}}\delta_h = 0.08$	ω	$2.17\ 10^{-6}\left(\frac{1}{cell\ day}\right)$	$tk\omega = 72.33$

that e.g. for the first equation that simulates the temporal change of c, the implementation uses Crank Nikolson for approximating c and its derivatives, grad and div, while it uses Forward Euler in order to approximate h, V, n and T. This holds similarly for the rest four equations of the non linear system.

V. Results

We have applied the PIHNA-LQ model for two cases. In the first case, the tumor is left to grow without treatment for 30 days, while in the second case radiotherapy is applied after 20 days of free growth for 10 more days. Table I depicts the parameters used in the model. The parameters used in the model were extracted from [12].

Before feeding the model with parameters, they were nondimensionalized according the formulas shown in the same table. The spatial nondimensionalization was built on a grid spanning an overall area of $L \times L$ for $L = 1cm$ and the temporal nondimensionalization was built on a time step of $\tau = 8h$. Thus, the new positions $\hat{x}$ are found according to $\hat{x} = \frac{x}{L}$ and the time is changed to $\hat{t} = \frac{t}{\tau}$. Lastly, the concentration c, h, n and v of normoxic, hypoxic, necrotic and endothelial cells, respectively, is nondimensionalized according to the capacity of cells $k = 10^8 \frac{cells}{cm^3}$, thus $\hat{c} = \frac{c}{k}, \hat{h} = \frac{h}{k}, \hat{n} = \frac{n}{k}$ and $\hat{v} = \frac{v}{k}$. Similarly, the maximum value of angiogenesis is altered according to the formula $\hat{a} = \frac{a}{a_{max}}$, where $a_{max} = 5.75\ 10^{-4} \frac{mmol}{cm^3}$.

For simulating radiotherapy, we set the LQ parameters as $R = 60Gy$, $A = 0.02Gy^{-1}$, $B = A/10$ and $OER = 3$. For both simulation cases, the initial states for all five cell populations was set at:

$$\hat{c}(\boldsymbol{x}, 0) = 0.93e^{-200(\boldsymbol{x}-\boldsymbol{x_0})^2}$$
$$\hat{h}(\boldsymbol{x}, 0) = 0$$
$$\hat{n}(\boldsymbol{x}, 0) = 0 \quad (6)$$
$$\hat{v}(\boldsymbol{x}, 0) = 0.01$$
$$\hat{a}(\boldsymbol{x}, 0) = 0$$

Row (a) of Fig. 2 presents the spatial concentration of normoxic, hypoxic and necrotic cells after 30 fictitious days of free glioma growth simulation. The computed graphs are depicted on a 40x40 grid (where $L = 1cm$ and $\boldsymbol{x_0} = (20,20)$). The row (b) depicts the respective graphs for the second case on the 30th day, assuming that radiotherapy treatment was applied for 10 days after 20 days of free growth (by applying the parameters of the previous section).

The results of our simulations computed that the peak value of the population of normoxic, hypoxic and necrotic cells is 0.7654, 0.2041 and 0.0139 in the case of free growth (PIHNA), while in the case of applying irradiation (PIHNA-LQ) these are 0.7614, 0.2906 and 0.8035 respectively. Moreover, the total viable cells $(c + h)$ on the 30th day decrease to $4.13k$ after therapy, compared to $5.23k$ for the case of free growth. Respectively, the total necrotic population (n) increases during radiotherapy, as expected.

These results indicate that the numerical framework presented can model the effect of radiotherapy by incorporating the LQ model into the PIHNA model. PIHNA-LQ model needs to be validated in a quantitative fashion against actual clinical data before and after radiotherapy treatment is applied to specific patients. We are currently in the final stage of glioma dataset collection.

VI. Conclusion

In this paper we presented an initial qualitative evaluation of a diffusion-reaction glioma growth model with normoxic, hypoxic and necrotic cells, incorporating vasculature, angiogenesis and radiotherapy treatment effect. We applied the Linear Quadratic model on the PIHNA model in two spatial dimensions.

The first results indicate that the radiotherapy can be simulated by this model.This is an important observation that needs further validation both from the experimental and the biological interpretation sides. Moreover, lysis of necrotic cells could be studied in future work by incorporating a subtraction term in the equation for necrotic cells.

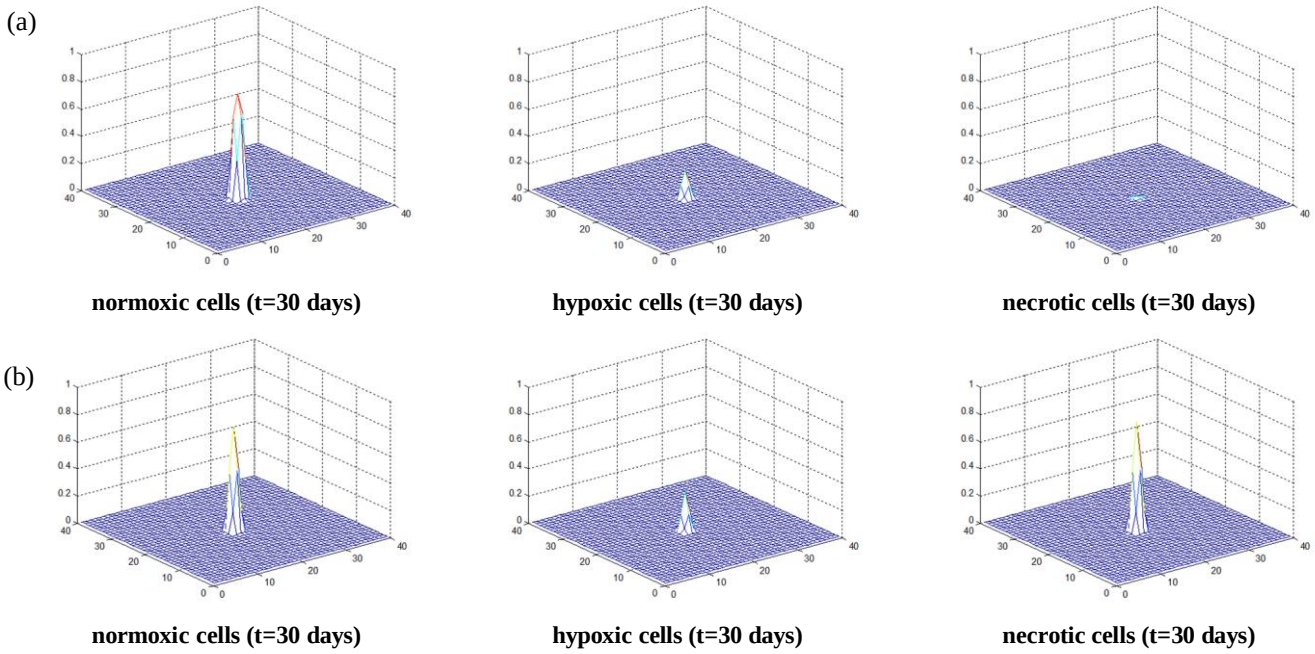


Fig. 2.The distribution of concentration for the three different populations of glioma cells (normoxic, hypoxic and necrotic cells) after applying the PIHNA model (a) and the PIHNA-LQ model (b). The graphs are extracted on the 30th day of simulation.

REFERENCES

[1] WHO, "http://www.who.int/mediacentre/factsheets/fs297," World Health Organization. Retrieved 2012-09-15.

[2] P. Boyle, and B. Levin, "World Cancer Report 2008,"IARC, 2010.

[3] Brain Tumors, "http://www.abta.org", American Brain Tumor Association. Retrieved 2012-09-15.

[4] A. Roniotis, K. Marias, V. Sakkalis, and M. Zervakis, "Diffusive modeling of glioma evolution: a review," *JBiomSciEng, ScientRes*, vol. 3, no. 5, pp.501-508,2010.

[5] A. Roniotis, G. Manikis, V. Sakkalis, M. Zervakis, I. Karatzanis, and K. Marias, "High grade glioma diffusive modeling using statistical tissue information and diffusion tensors extracted from atlases," *IEEE Trans Inform. Tech. in Biom*, vol. 16, no. 2, pp. 255-263, 2012.

[6] A. Roniotis, V.Sakkalis, I.Karatzanis, M. E. Zervakis, and K. Marias,"In-depth analysis and evaluation of diffusive glioma models,"*IEEE Trans Inform. Tech. in Biomedicine*, vol. 16, no. 3, pp. 299-307, 2012.

[7] K. R. Swanson, E. C. Alvord, and J. D. Murray, "A Quantitative Model for Differential Motility of Gliomas in Grey and White Matter," *Cell Proliferation*, vol. 33, no. 5, pp. 317-330, 2000.

[8] S. Jbabdi, E. Mandonnet, and H. Duffau, "Simulation of anisotropic growth of low-grade gliomas using diffusion tensor imaging," *MagnReson Med*, vol. 54, pp. 616-24, 2005.

[9] O. Clatz, M. Sermesant, P. Bondiau, H. Delingette, S. K. Warfield, G. Malandain, and N. Ayache, "Realistic simulation of the 3-D growth of brain tumours in MR images coupling diffusion with biomechanical deformation", *IEEE Transactions on Medical Imaging*, vol. 24, pp. 1334-1346, 2005.

[10] G. Stamatakos, V. Antipas, N. Uzunoglu, and R. Dale, "A four-dimensional computer simulation model of the in vivo response to radiotherapy of glioblastomamultiforme: studies on the effect of clonogenic cell density,"*Br J Radiol*, vol. 79, pp. 389-400, 2006.

[11] G. Stamatakos, V. Antipas, N. Ozunoglu, "A patient-specific in vivo tumor and normal tissue model for prediction of the response to P. Hinow, P. Gerlee, L. J. McCawley, V. Quaranta, M. Ciobanu, S. Wang, J. M. Graham, B. P. Ayati, J. Claridge, K. R. Swanson, M. Loveless, and A. R., Anderson, "A spatial model of tumor-host interaction: application of chemotherapy," Math BiosciEng,vol. 6, no. 3, pp. 521-46, 2009.

[12] K. R. Swanson, R. C. Rockne, J. Claridge, M. A. Chaplain, E. C. Alvord Jr, and A. Anderson, "Quantifying the role of angiogenesis in malignant progression of gliomas: In Silico modeling integrates imaging and histology," *Cancer Res*, vol. 71, pp. 7366-7375, 2011

[13] G. Steel, "Basic Clinical Radiobiology, 3rd Edition," Arnold, London, 2002.

[14] E. Hall, "Radiobiology for the Radiologist,"*J B Lippincott*, Philadelphia, 1994.

[15] R. Rockne, J. K. Rockhill, M. Mrugala, A. M. Spence, I. Kalet, K. Hendrickson, A. Lai, T. Cloughesy, E. C. Alvord, and K. Swanson, "Predicting the efficacy of radiotherapy in individual glioblastoma patients in vivo: a mathematical modeling approach," *Phy Med Biol*, vol. 55, pp. 3271-3285, 2010.

[16] A. Roniotis, K. Marias, V. Sakkalis, G. C. Manikis, and M. Zervakis, "Simulating radiotherapy effect in high grade glioma by using diffusive modeling and brain atlases," *J Biomed Biotech* (in press), 2012

[17] L. Harrison, M. Chadha, R. J. Hill, K. Hu, and D. Shasha, "Impact of tumor hypoxia and anemia on radiation therapy outcomes," *Oncologist* vol. 7, no. 6, pp. 492–508, 2002.

[18] J. Brown, and W. Wilson, "Exploiting tumour hypoxia in cancer treatment", *Nat Rev Cancer*, vol. 4, no. 6, pp. 437-47, 2004

Modeling Nephroblastoma Treatment Response Cases with *In-Silico* Scenarios*

Eleni Ch. Georgiadi, Dimitra D. Dionysiou, Norbert Graf and Georgios S. Stamatakos, *Member, IEEE*

***Abstract*— Two blastemal nephroblastoma cases have been successfully clinically adapted under two simulation scenarios of a clinically-oriented multiscale computational model, providing insight into the tumor characteristics.**

I. Introduction

In the last years many computational models have been produced [1-3] in order to simulate the hypercomplex phenomenon of cancer. A predominately discrete, clinical trial-driven, 4D simulation model of Wilms' tumor, simulating free growth and response to preoperative chemotherapy has been developed by the In Silico Oncology Group-National Technical University of Athens (ISOG-NTUA) [4-5]. The simulation approach is "top-down" integrating multiple levels of biological complexity. A level of confidence of the model has been already set [5]. In the framework of further adaptation and validation of the model, two unilateral nephroblastoma cases treated with preoperative chemotherapy are simulated under two virtual scenarios. Successful adaptation has been achieved and valuable hints concerning tumor characteristics are revealed.

II. In Silico Model Description

A. Modeling tumor initialization and advancement

The model structure and algorithm has been analytically presented in previous publications [4,5] and will be shortly described here:

Available patient-individualized imaging data are used to define the area of interest which is discretized by a 3D cubic grid. The "voxel" of the grid is called ''geometrical cell'' (GC). Each GC within the tumor volume is occupied and initialized with a population of cells (e.g.10^9 cells/cm^3) [6],

The cell kinetics of tumor growth and response to chemotherapy incorporated in the model is illustrated in figure 1. The basic cell categories are the following: proliferating, dormant, dead and differentiated cells. Proliferating cells include stem (cells of unlimited mitotic potential) and LIMP cells (LImited Mitotic Potential cells) lying in G1, S, G2 or M cell cycle phases. Dormant cells include stem and LIMP cells residing in dormant state, whereas dead cells include necrotic and apoptotic cells. Differentiated cells have no mitotic potential. The transitions between the different states indicated in figure 1 incorporate several biological phenomena.

At each computational time step the grid is scanned in two phases: firstly the rules of the cytokinetic model (figure 1) are applied and secondly a geometrical reconstruction of the tumor (shrinkage/expansion) is performed.

The tumor dynamics model parameters are defined in the appendix (Table III). Plausible reference values of the various model parameters retrieved from literature in conjunction with accumulative basic science and clinical experience are noted in Table I.

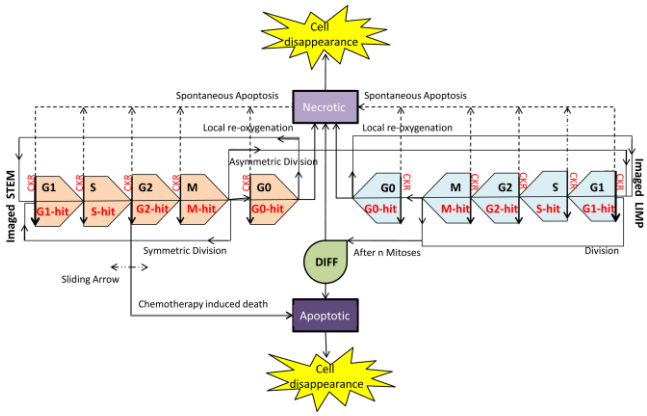
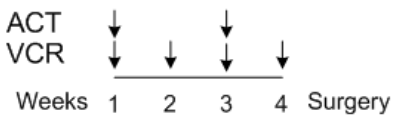


Figure 1. Tumor cell growth and therapy kinetics



ACT: Actinomycin-D. 45μg/kg i.v. bolus injection (max 2000μg)
VCR: Vincristine. 1.5mg/m^2 i.v. bolus injection (max 2.0mg)

If body weight < 12kg: dose reduction to 2/3 for each drug
Major intolerance: doses on the next course should be reduced to 2/3

Figure 2. The simulated Wilms tumor chemotherapy treatment protocol (SIOP/GPOH clinical trial)

* This work has been supported in part by the European Commission under the projects ACGT ''Advancing Clinicogenomic Trials on Cancer'' (FP6-2005-IST-026996), TUMOR (Transatlantic Tumor Model Repositories) (FP7-ICT-2009.5.4-247754) and p-Medicine (From data sharing and integration via VPH models to personalized medicine) (FP7-ICT-2009.5.3-270089).

E. Ch. Georgiadi (e-mail: egeorg@central.ntua.gr), Dimitra D. Dionysiou (e-mail: dimdio@esd.ece.ntua.gr) and Georgios S. Stamatakos (corresponding author: tel.: +30 2107722287 ; fax: +30 2107723557; e-mail: gestam@central.ntua.gr) are with the In Silico Oncology Group, Institute of Communication and Computer Systems, School of Electrical and Computer Engineering, National Technical University of Athens.

N. Graf (e-mail: Norbert.Graf@uniklinikum-saarland.de) is with Universität des Saarlandes, Klinik für Pädiatrische Onkologie und Hämatologie, Homburg, Germany

TABLE I. TUMOR DYNAMICS MODEL PARAMETERS

Symbol (units)	Reference Value	References	T1[a]	T2[b]	T3[c]	T4[d]
Tc (h)	23.0	[21]	23.0	40	23.0	55
T_{G0} (h)	96	[22]	96	96	96	40
T_N (h)	20	[23,24,25]	20	20	20	120
T_A (h)	6	[42, 43]	6	6	6	6
R_A (h^{-1})	0.001	Derived from T_A, based on [26,27]	0.001	0.0008	0.001	0.001
R_{ADiff} (h^{-1})	0.003		0.003	0.003	0.003	0.05
R_{NDiff} (h^{-1})	0.001	Derived from T_N, based on [23,25]	0.001	0.001	0.001	0.05
P_{G0toG1}	0.01		0.01	0.01	0.01	0.01
N_{LIMP}	3		3	3	3	3
P_{sym}	0.45		0.71	0.44	0.45	0.465
P_{sleep}	0.28		0.40	0.28	0.28	0.36
CKR_{VCR}	0.3	Derived based on [7,8]	0.32	0.3	0.252	0.33
CKR_{ACT}	0.2	Derived based on [9,10]	0.213	0.2	0.168	0.22
CKR_{TOTAL}*	0.5	Additive drug effect considered	0.533	0.5	0.42	0.55

a. T1: Case A-Scenario A
b. T2: Case A-Scenario B
c. T3: Case B-Scenario A
d. T4: Case B-Scenario B

* CKR_{TOTAL} is not an independent parameter of the model

B. *Modeling tumor chemotherapy*

Chemotherapeutically induced death is incorporated in the model as well. Pharmacokinetics and pharmacodynamics of the chemotherapeutic drugs used, [8-10] determine the fraction of cells which enter a rudimentary cell cycle-called hit-(Figure 1) leading to apoptotic death. The actual time of death is also defined.

A protocol of preoperative chemotherapy with a combination of actinomycin-D and vincristine for unilateral stage I-III nephroblastoma tumors, treated according to the SIOP 2001/GPOH clinical trial (Figure 2), has been specifically simulated.

According to the SIOP 2001/GPOH clinical trial protocol, vincristine i.v. bolus injection is directly followed by an i.v. bolus injection of actinomycin-D, with no delay in-between.

In the simulation model vincristine is assumed to bind at cells at all cycling phases and lead to apoptosis at the end of M phase [11-13]. Actinomycin-D is considered to bind to cells at all phases (including G0) and lead to apoptosis at the end of the S phase [14].

The estimation of typical values of the cell kill ratios of vincristine and actinomysin-D is based on relevant pharmacokinetics and pharmacodynamics literature [14, 7-10].

An additive drug effect of vincristine and actinomycin-D has been assumed for all active cell cycle phases. For dormant cells only actinomycin-D exerts a cytotoxic effect.

III. NEPHROBLASTOMA CASES

Several nephroblastoma tumor cases collected in the context of SIOP 2001/GPOH trial have been modeled so far. The in silico response of two (case A and case B) nephroblastoma cases of stromal type to preoperative combined chemotherapy with actinomycin-D and vincristine is presented.

The anonymized imaging and clinical data have been provided by the clinicians. The initial and final virtual tumors have been spatiotemporally initialized based on MRI images of the clinical tumors collected at two time instants before the start of chemotherapy (2 days for case A and 6 days for case B) and after the completion (4 days for case A and 0 days for case B). The chemotherapeutic scheme administered is described in figure 2.

Two adaptation scenarios are considered. The following assumptions have been made based on literature and imaging clinical information:

a. Doubling time in the range of 11–40 days [15]–[19]
b. Growth fraction 1.5-20% for stromal type nephroblastomas [20]
c. The imaging data-specified volume reduction of case A is 65% and of case B is 32%.

As a first adaptation scenario (A), the two nephroblastoma tumors are considered to have common growth kinetics features as they are of the same histological type (blastemal). The cell kill ratio of the chemotherapeutic drugs is adapted in order to simulate the volume reduction induced by chemotherapy.

As a second adaptation scenario (B), the effect of the chemotherapeutic drugs is considered common for both cases and the growth rate of the tumors is adapted within the range defined by reported literature, by perturbing the value of an important growth kinetic parameter according to model sensitivity analyses [4], the fraction of stem cells that perform symmetric division (Psym).

The above assumptions in conjunction with accumulated basic science and clinical experience-plausible values set the tumor dynamic model parameters (Table I) for cases A, B and under scenarios A, B.

The time course of the volume of the four virtual tumors is presented in figure 3. The resultant virtual tumor characteristics and volume reduction are given in table II.

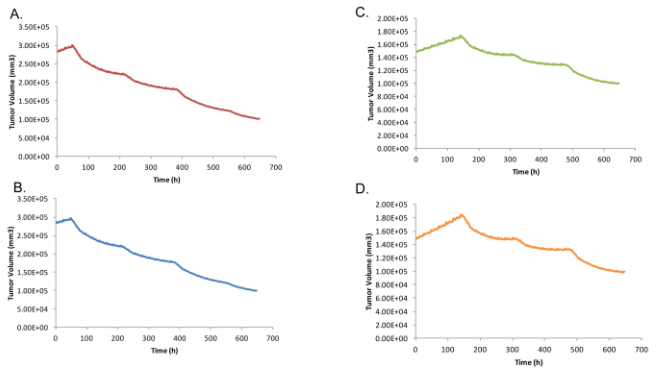


Figure 3. Tumor volume time evolution for the four simulation scenarios (Table 1): A: T1, B: T2, C: T3, D: T4

TABLE II. TUMOR CHARACTERISTICS AND VOLUME REDUCTION PERCENTAGES FOR THE FOUR VIRTUAL TUMOR SCENARIOS (TABLE 1)

Resultant initial tumor characteristics	T1[a]	T2[b]	T3[c]	T4[d]
Volume Doubling Time, $T_d = \ln 2/k$ (days)	28.5	40.37	28.5	19.8
Initial percentage of proliferating cells (Growth Fraction) (%)	13.38	12.62	13.38	14.52
Tumor volume reduction percentage (%)	64.24	65.02	32.17	32.58

a. T1: Case A-Scenario A
b. T2: Case A-Scenario B
c. T3: Case B-Scenario A
d. T4: Case B-Scenario B

IV. DISCUSSION-CONCLUSION

Two blastemal type nephroblastomas (case A and B) are treated according to SIOP 2001/GPOH clinical trial protocol for uniratelar stage I-III nephroblastoma tumors. The specified by the available imaging data (pre and post) chemotherapy induced reduction of the tumor volume is 65% and 32% respectively.

The difference in the chemotherapy outcome might reflect different tumor dynamics (Scenario B) or different response to the chemotherapeutic treatment (Scenario A).

These two scenarios have been simulated and adapted to literature and imaging clinical information. The imaging data-specified induced volume reduction for both cases has been successfully adapted. The growth fraction and the doubling time for both cases stand within the plausible range.

Under Scenario A, both tumors have the same doubling time and growth fraction. In order to achieve the specified clinical volume reduction, the additive drug effect relative difference (to the reference value) between the two cases is 23%.

Under Scenario B, the effect of therapy is the same for both tumors and the growth kinetics of tumors differ. Case A which presents greater volume reduction, appears to have higher doubling time and lower growth fraction. Conclusively the less aggressive tumor reveals better response to chemotherapy (Table II).

The fitting of the selected nephroblastoma cases to the clinical data serves as an example, demonstrating a possible procedure towards the formation of virtual scenarios, based on the combined use of the available clinical and literature data. As indicated by the previous adaptation examples, adequate "tuning" of the simulation results could give valuable hints concerning tumor characteristics for which actual estimations might be missing in each case considered.

APPENDIX

TABLE III. DEFINITION OF TUMOR DYNAMICS MODEL'S PARAMETERS

Symbol (units)	Definition
Tc (h)	Cell cycle duration
T_{G0} (h)	G0 (dormant phase) duration, i.e. time interval before a dormant cell dies through necrosis
T_N (h)	Time needed for necrosis to be completed and its lysis products to be eliminated from the tumor
T_A (h)	Time needed for apoptosis to be completed and its products to be eliminated from the tumor
R_A (h^{-1})	Apoptosis rate of living stem and LIMP tumor cells (fraction of non-differentiated cells dying through apoptosis per hour)
R_{ADiff} (h^{-1})	Apoptosis rate of differentiated tumor cells per hour
R_{NDiff} (h^{-1})	Necrosis rate of differentiated tumor cells per hour
P_{G0toG1}	The fraction of stem or LIMP cells having just left the G0 compartment that re-enter the cell cycle
N_{LIMP}	The maximum number of mitoses that a LIMP cell can perform before becoming terminally differentiated
P_{sym}	Fraction of stem cells that perform symmetric division.
P_{sleep}	Fraction of cells that enter G0 phase following mitosis
CKR_{VCR}	Cell kill ratio for the specific vincristine dose
CKR_{ACT}	Cell kill ratio for the specific actinomycin-D dose
CKR_{TOTAL}*	Combined cell kill ratio of the two drugs (dependent parameter)

REFERENCES

[1] Anderson AR, Quaranta V (2008) Integrative mathematical oncology. Nat Rev Cancer 8: 227–234.

[2] Deisboeck TS, Zhang L, Yoon J, Costa J (2009) *In silico* cancer modeling: is it ready for prime time? Nat Clin Pract Oncol 6(1): 34–42.

[3] Multiscale Cancer Modeling, T. Deisboeck and G. Stamatakos, Eds. Boca Raton, Florida, USA: CRC Press, 2010

[4] G.S.Stamatakos, E.C.Georgiadi, N.Graf, E.A.Kolokotroni, D.D.Dionysiou, Exploiting clinical trial datadrastically narrows the window of possible solutions to the problem of clinical adaptation of a multiscale cancer model, Plos One 6(3), e17594, 2011.

[5] Georgiadi EC, Stamatakos GS, Graf NM, Kolokotroni EA, Dionysiou DD, et al. (2008) Multilevel Cancer Modeling in the Clinical Environment: Simulating the Behavior of Wilms Tumor in the Context of the SIOP 2001/GPOH Clinical Trial and the ACGT Project.

[6] Castorina P, Carcò D, Guiot C, Deisboeck TS (2009) Tumor Growth Instability and Its Implications for Chemotherapy. Cancer Res 69(21): 8507–15.

[7] Groninger E, Meeuwsen-de Boer T, Koopmans P, Uges D, Sluiter W, et al. (2002) Pharmacokinetics of Vincristine Monotherapy in Childhood Acute Lymphoblastic Leukemia. Pediatric Research 52: 113–118.

[8] Dahl WN, Oftebro R, Pettersen EO, Brustad T (1976) Inhibitory and cytotoxic effects of Oncovin (Vincristine Sulfate) on cells of human line NHIK 3025. Cancer Res 36: 3101–3105.

[9] Sawada K, Noda K, Nakajima H, Shimbara N, Furuichi Y, et al. (2005) Differential cytotoxicity of anticancer agents in pre- and post-immortal lymphoblastoid cell lines. Biol Pharm Bull 28: 1202–1207.

[10] Veal GJ, Cole M, Errington J, Parry A, Hale J, et al. (2005) Pharmacokinetics of Dactinomycin in a pediatric patient population: a United Kingdom Children's Cancer Study group study. Clin Cancer Res 11(16): 5893–5899.

[11] Beck WT, Cass CE, Houghton PJ (2000) Microtubule-targeting anticancer drugs derived from plants and microbes: Vinca alkaloids, taxanes and epothilones. In: Bast RC, Kufe DW, Pollock RE, Weichelbaum RR, Holland JF, Frei E, editors. Cancer Medicine 5th ed. Hamilton , Ontario, Canada: BC Decker.
[12] Wood KW, Cornwell WD, Jackson JR (2001) Past and future of the mitotic spindle as an oncology target. Curr Opin Pharmacol 1(4): 370–377
[13] Pinkerton CR, McDermott B, Philip T, Biron P, Ardiet C, et al. (1988) Continuous vincristine infusion as part of a high dose chemoradiotherapy regimen: drug kinetics and toxicity. Cancer Chemother Pharmacol 22: 271–274
[14] Salmon SE, Sartorelli AC (2001) Cancer Chemotherapy. In: Katzung BG, editor. Basic & Clinical Pharmacology. International Edition: Lange Medical Books/McGraw-Hill. pp. 923–1044.
[15] Tan TY, Amor DJ (2006) Tumour surveillance in Beckwith-Wiedermann syndrome and hemihyperplasia: A critical review of the evidence and suggested guidelines for local practice. Journal of Paedeatrics and Child Health 42: 486–490.
[16] Shackney SE, McCormack GW, Cuchural GJ (1978) Growth rate patterns of solid tumours and their relation to responsiveness to therapy. Ann Intern Med 89: 107–121.
[17] Craft AW (1999) Growth rate of Wilms' tumour. The Lancet 354(9184): 1127
[18] Zoubek A, Slavc I, Mann G, Trittenwein G, Gadner H (1999) Natural course of a Wilms' tumour. Lancet 354: 344.
[19] Carré A, Frantz CN, Weksberg R, Nicholson L, Ciarlo L, et al. (2005) Wilms tumor in an 11-year old with hemihyperplasia. Am J Med Genetics 139A: 165–166.
[20] Berrebi D, Leclerc J, Schleiermacher G, Zaccaria I, Boccon-Gibod L, et al. (2008) High cyclin-E staining index in blastemal, stromal or epithelial cells is correlated with tumor aggressiveness in patients with nephroblastoma. PLoS ONE 3(5): e2216.
[21] Revazova ES, Petrova AS (1981) Cell cycle and proliferative pool of human tumor strains transplanted into athymic mice. Biull Eksp Biol Med 92: 335–337.
[22] Maseide K, Rofstad EK (2000) Mathematical modeling of chronical hypoxia in tumors considering potential doubling time and hypoxic cell lifetime. Radiother Oncol 54: 171–177.
[23] Duechting W, Ulmer W, Lehrig R, Ginsberg T, Dedeleit E (1992) Computer simulation and modeling of tumor spheroids growth and their relevance to optimization of fractionated radiotherapy. Strahlenther Onkol 168(6): 354–360.
[24] Titz B, Jeraj R (2008) An imaging-based tumour growth and treatment response model: investigating the effect of tumour oxygenation on radiation therapy response. Phys Med Biol 53: 4471–4488.
[25] Wein LM, Cohen JE, Wu JT (2000) Dynamic optimization of a linear-quadratic model with incomplete repair and volume-dependent sensitivity and repopulation. Int J Radiat Oncol Biol Phys 47(4): 1073–1083.
[26] Ribba B, Colin T, Schnell S (2006) A multiscale mathematical model of cancer, and its use in analyzing irradiation therapies. Theor Biol Med Model 3: 7.
[27] Dewey W, Ling CC, Meyn RE (1995) Radiation-induced apoptosis: relevance to radiotherapy. Int J Radiat Oncol Biol Phys 33(4): 781–796.

Clinical Evaluation of DoctorEye Platform in Nephroblastoma*

Ruslan David, Norbert Graf, Ioannis Karatzanis, Holger Stenzhorn, Georgios C. Manikis, Vangelis Sakkalis, Georgios S. Stamatakos, *Member, IEEE*, Konstantinos Marias, *Member, IEEE*

Nephroblastoma is the most common malignant renal tumor in children. Today about 90 % of the patients can be cured by chemotherapy and surgery. Most of the patients are enrolled in prospective clinical trials. In the SIOP (International Society of Pediatric Oncology) approach all children do receive preoperative chemotherapy to shrink the tumor before surgery. Patients with an excellent response to preoperative chemotherapy have a better outcome than those with a poor response. Response is not only defined by tumor volume shrinkage but also by the vanishing of blastemal tumor cells. Imaging studies are used to calculate tumor volume and to define vital, necrotic and cystic areas within a tumor. In case of nephroblastoma it would be of most importance to further separate the vital tumor in blastemal and non-blastemal components. This would allow changing treatment at an early phase in case of remaining blastema. Using the DoctorEye software, tumors can be easily rendered and histograms of the signal intensities within a tumor are possible to calculate. Our preliminary results show that these histograms give further insight in the tumors in single patients by correlating them with the histological findings.

I. Introduction

Nephroblastoma is the most common malignant renal tumor in children. Dramatic improvements in survival have occurred as the result of advances in interdisciplinary treatments. They are based on several multicenter trials and studies conducted by the SIOP in Europe and COG (Children's Oncology Group) in North America. Main objectives of these trials and studies are to treat patients according to well-defined risk groups in order to achieve highest cure rates, to decrease the frequency and intensity of acute and late toxicity and to minimize the cost of therapy. In that way the SIOP trials and studies largely focus on preoperative therapy [1]. The concept of neoadjuvant chemotherapy plays an important role in the treatment for most pediatric solid tumors today [2]. The complete surgical removal of a shrunken tumor is facilitated, mutilation caused by surgical procedures is minimized or avoided and micrometastases, not visible at diagnosis, are treated as early as possible. Besides that, response to treatment can be measured individually by tumor volume reduction and / or percentage of therapy induced necrosis at the time of surgery in the histological specimen. This gives an early individual prognostic parameter and is used for individualizing postoperative treatment [1].

DoctorEye [3] is an open access, flexible and easy to use platform, for intuitive annotation and segmentation of tumor regions. Its clinically driven development followed an open modular architecture focusing on plug-in components. DoctorEye's main advantage is that the user can quickly and accurately delineate complex areas in medical images in contrast with other platforms that do not facilitate the delineation of areas with complicated shapes. Additionally, multiple labels can be set to allow the user to annotate and manage many different areas of interest in each selected slide. The close collaboration with clinicians in designing the platform has ensured that it can be effectively used in the clinical setting.

Another reported feature [4] that adds value to the platform is that it allows computational "in-silico" models of cancer growth and simulation of therapy response to be easily plugged in, in order to provide a future integrated platform for modeling assisted therapy decision making. Currently, DoctorEye's development team is working towards incorporating such models in the platform. In this context, the platform could also serve as a validation environment where the simulation predictions could be compared with the actual therapy outcome in order to achieve a global optimization of the modeling modules. DoctorEye platform has been actively developed in the frames of ContraCancrum European Commission ICT research project [5] and at the moment it serves as an intuitive 3D annotation system.

II. Procedures and Methods

A. DoctorEye platform design

DoctorEye platform is based on the .NET framework architecture and can be used in any Windows-based computer. The graphical interface is based on well-known Microsoft Office applications to ensure a user-friendly environment.

We have been using DoctorEye over a large amount of datasets and found the platform's environment simple, intuitive and user friendly. Only some basic knowledge and experience of annotation/simulation processes are required to use the system.

B. Segmentation Methods Applied

Segmentation is the process dividing an image into regions with similar properties such as gray level, color, texture, brightness, and contrast. The role of segmentation is to subdivide the objects in an image; in case of medical image segmentation the aim is to:

*Research supported by the European Commission under the project p-medicine (FP7-ICT-2009.5.3, Grant agreement number 270089).

N. G. Author is with the Saarland University, Dep. Pediatric Oncology and Hematology, 66421 Homburg, Germany (phone: 0049 6841 1628397; fax: 0049 6841 1628302; e-mail: graf@uks.eu).

R. D. and H. S. are with the Saarland University, Dep. Pediatric Oncology and Hematology, 66421 Homburg, Germany (e-mails: ruslan.david@uks.eu; holger.stenzhorn@uks.eu).

I. K., G. C. M., V. S. and K. M. are with Institute of Computer Science at FORTH, Vassilika Vouton, GR-70013 Heraklion, Crete, Greece (e-mails: karatzan@ics.forth.gr; gmanikis@gmail.com; sakkalis@ics.forth.gr; kmarias@ics.forth.gr).

G. S. S. is with Institute of Communications and Computer Systems, National Technical University of Athens, *In Silico* Oncology Group, Iroon Polytechniou 9, Zografos, Greece (email: gestam@central.ntua.gr).

- Study anatomical structure
- Identify Region of Interest i.e. locate tumor, lesion and other abnormalities
- Measure tissue volume to measure growth of tumor (also decrease in size of tumor with treatment)
- Help in treatment planning prior to radiation therapy; in radiation dose calculation

Automatic segmentation of medical images is a difficult task as medical images are complex in nature and rarely have any simple linear feature.

Medical image segmentation has been a subject of the advanced research activities in the past years. Both the automatic and semiautomatic detection and delineation of tumors, generally involve different types of tissue and fuzzy boundaries rendering them difficult to segment. At this phase of the development, we focus mostly on MR imaging modality as the clinicians in Wilms tumor cases have mostly used MRI.

Two different, in terms of underlying theoretical concepts, segmentation algorithms are currently available in the Doctor Eye platform: namely the "Magic Wand" and "Spatially Adaptive Active Contours", additionally, due to program flexibility any other algorithm can be easily plugged-in. The first is based on image intensity, whereas the "active contours" are model-based.

III. Segmentation of Wilms' Tumor

Tumor images segmentation activities in general are very complex due to high variability of tumor structures. Since manual segmentation is time-consuming, and, on the other side, a fully automatic technique is not applicable because the enrolment of the clinician's expertise/supervision is required, the best choice seems to be the semi-automatic technique. In order to assure a knowledge-based background for further research activities related to semi-automatic technique, in the frames of our research activities the manual segmentation/annotation of the Wilms' tumor has been performed.

Our preliminary results could be divided in two conventional parts where the first part presents the segmentation/annotation of Wilms' tumor by using the semi-automatic technique, and the second part presents the segmentation/annotation of Wilms' tumor performed by using the manual technique:

A. *Segmentation of Wilms' Tumor (semi-automatic technique)*

Segmentation of Wilms' Tumor by using a semi-automatic snake-based approach for tumor segmentation, based on the use of locally adaptable parameters has been described in details by Farmaki et al. [6]. This reported method exploits gradient and curvature characteristics to compute adaptive sets of parameters which guarantee that the evolving snake is rigid and forceful inside the desired boundary, in order to overcome small inhomogeneities, and very flexible near the tumor borders, in order to accurately adjust to boundary details.

It clearly improved tumor segmentation results comparing to previous methods, since the average overlap, produced by the described method's application on several medical volumes, was 89%, while the corresponding overlap for traditional snakes and region growing was 82.5 and 59.2%, respectively.

The figure bellow (Figure 1) [6] presents the results of the application of the descried method on three MR scans, here the spatially adaptive active contours was used. (a) shows the clinician's annotation of the tumor, (b) shows the traditional active contour's result, and (c) depicts the result obtained.

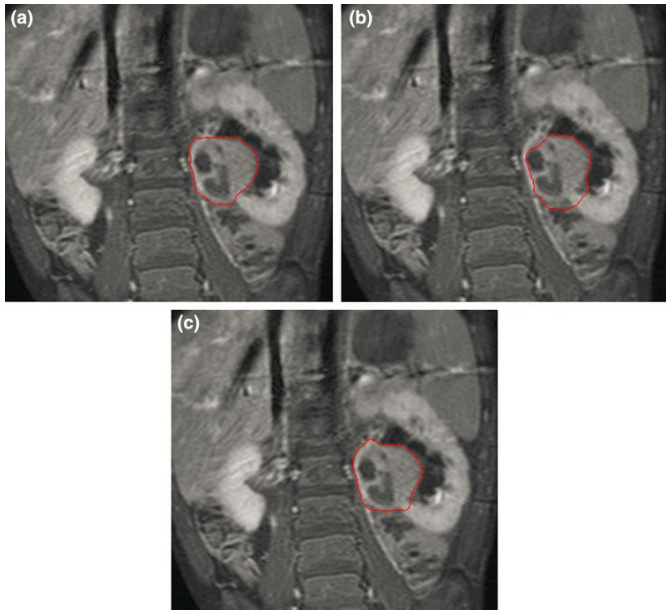


Figure 1. Tumour segmentation on MR abdominal images: (a) clinician's annotation, (b) result using traditional snakes, (c) result using spatially adaptive active contours.

B. *Segmentation of Wilms' Tumour (manual technique)*

The manual segmentation of the Wilms' tumour technique has been performed in our research to annotate precisely and with minimal possible errors the tissue of Wilms' tumour. This approach, at first stage, allowed us to benchmark the manual and semi-automatic techniques. Also, it allowed us to assess the output of the DoctorEye' segmentations tools: Pencil, Magic Wand, Active Contours (Greedy & Snake).

DoctorEye platform has advanced segmentation/ annotation tools which allow the end-user to select with high accuracy the Wilms' tumor tissue but according to our preliminary results the most precise seems to be the "Pencil" interface. The only disadvantages are the time related restrictions; segmentation/annotation of the areas of interest with the "Pencil" tool requires much more time comparing to Magic Wand and Active Contours (Greedy & Snake) tools. Nevertheless, the results of "Pencil" tool annotation/segmentation are precise and with minimal possible errors. An example output (in red color) of the "Pencil" segmentation/annotation is presented in Figure 2.

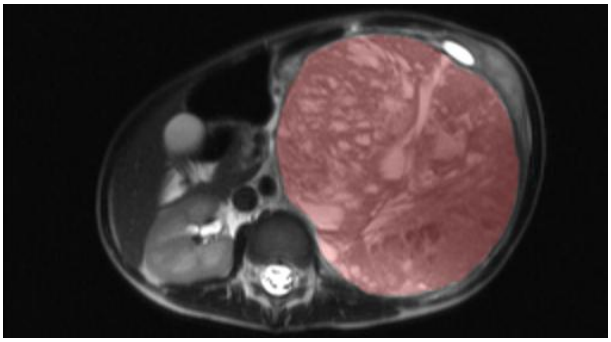


Figure 2. Pencil segmentation/annotation tool example

IV. Addressing pre/post Chemotherapy Changes

The main goal of the conducted Wilms' tumor segmentation activities was to address the pre/post chemotherapy changes in MR images. The aim of this research was to identify and to report changes after chemotherapy in tumor volume estimated with DoctorEye platform, and to compare the histopathological class to these changes.

DoctorEye platform allows the end-users to generate the histograms of the segmented/annotated region(s). A histogram could be generated for one image or for all images in the available data set. An image histogram is a graphical representation of the tonal distribution in a digital image. The horizontal axis of the graph represents the tonal variations, while the vertical axis represents the number of pixels in that particular tone. The left side of the horizontal axis represents the black and dark areas, the middle represents medium grey and the right hand side represents light and pure white areas.

Wilms' tumor regions in MR pre/post chemotherapy images of 20 cases/patients have been segmented/annotated. From 16 cases the following criteria were available for further analysis (Table 1):

- confirmed unilateral Wilms' tumor (post-surgery histopathology results)
- presence of pre/post chemotherapy DICOM images (MRI) data sets
- enrolment in SIOP 2001 trial (signed informed consent)
- availability of Wilms' tumor's histopathology type data (post-surgery histopathology)

Close to annotated/segmented MR images, DoctorEye platform has been used as well to identify the pre/post chemotherapy tumor volume change. The generated histograms of the annotated/segmented MR images have been divided in pre/post chemotherapy charts (Figures 3, 4, 5).

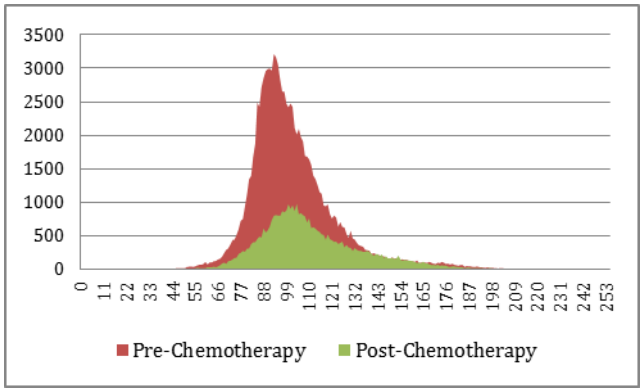


Figure 3. Case 4 (C4) - "Median" Post-chemotherapy histogram

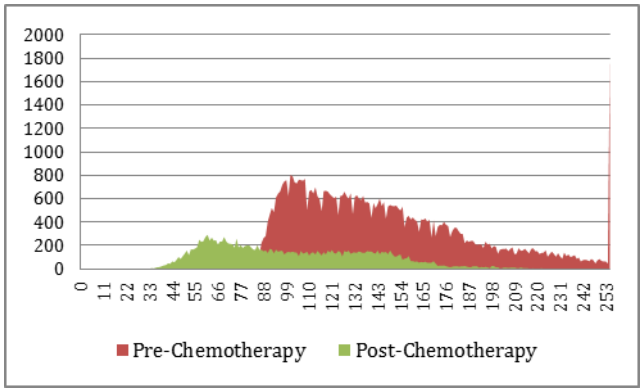


Figure 4. Case 7 (C7) - "Left shift" Post-chemotherapy histogram

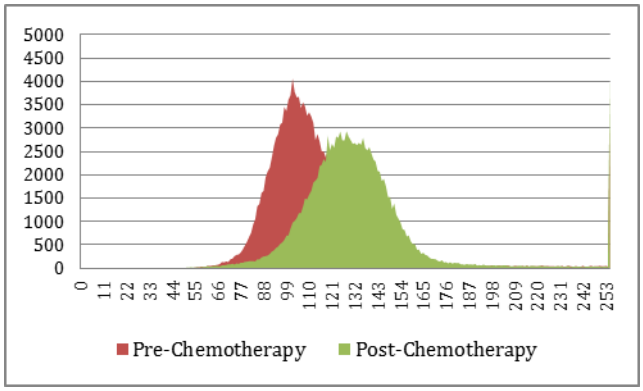


Figure 5. Case 10 (C10) - "Right shift" Post-chemotherapy histogram

Wilms' tumor presents as a large, solid tumor of renal origin. The tumor may be homogeneous, but typically appears heterogeneous with intermediate signal intensity on T1-weighted images and high signal intensity on T2-weighted images [7]. In our cases we decided to annotate/segment only T2-weighted images due to increased signal intensity specifically in T2 images [8]. Size and extent of tumor regions have been precisely identified and annotated/segmented.

TABLE I. THE RESULTS OF WILMS' TUMOUR SEGMENTATION (16 CASES)

Case (C)	Necrosis / Regression	Histopathological type	Post chemotherapy histogram	Tumor Volume (ml)	
				Pre chemotherapy	*Post chemotherapy*
C1	80%	70% Blastema 30% Epithelia 0% Stroma	Left shift	443,075	182,873
C2	80%	80% Blastema 20% Epithelia 0% Stroma	Left shift	76,514	7,025
C3	20%	50% Blastema 25% Epithelia 25% Stroma	Left shift	143,629	19,385
C4	5%	0% Blastema 100% Epithelia 0% Stroma	Median	288,629	86,017
C5	2%	4% Blastema 2% Epithelia 94% Stroma	Slight left shift	592,262	497,733
C6	50%	10% Blastema 60% Epithelia 30% Stroma	Slight left shift	967,886	97,465
C7	10%	20% Blastema 30% Epithelia 50% Stroma	Left shift	1041,203	390,448
C8	10%	0% Blastema 50% Epithelia 50% Stroma	Almost median	319,142	144,28
C9	30%	40% Blastema 30% Epithelia 30% Stroma	Right shift	1117,213	1292,07
C10	5%	50% Blastema 45% Epithelia 5% Stroma	Right shift	804,95	756,57
C11	95%	90% Blastema 10% Epithelia 0% Stroma	Left shift	748,875	109,334
C12	30%	2% Blastema 8% Epithelia 90% Stroma	Almost median	964,169	515,963
C13	90%	50% Blastema 50% Epithelia 0% Stroma	Slight left shift	105,884	11,504
C14	70%	90% Blastema 0% Epithelia 10% Stroma	Right shift	1691,756	974,453
C15	25%	10% Blastema 60% Epithelia 30% Stroma	Right shift	800,874	756,212
C16	30%	15% Blastema 25% Epithelia 60% Stroma	Left shift	895,801	615,809

V. PRELIMINARY FINDINGS AND CONCLUSIONS OF FURTHER WORK

DoctorEye's manual and semi-automatic segmentation techniques combined with integrated correction tools assist in the fast and refined delineation of Wilms' tumors and different users are able to add different components such as tumor growth and simulation algorithms for improving therapy planning. DoctorEye platform has been successfully tested and practically used over a large number of Wilms' tumor MRI datasets and it ensured stability, usability, extensibility and robustness with promising results.

Our preliminarily findings confirm the evidence of significant shrinkage or volume change of the tumors following primary chemotherapy (excepting the case C9). Changes in Wilms' tumor after chemotherapy have been addressed for the first time (to our knowledge) by using the advanced segmentation/annotation techniques and the histogram generation interfaces of DoctorEye platform. This research was the first report to specifically present pre/post chemotherapy changes of Wilms' tumor histograms close to the related volume changes.

We hypothesize that the histograms of Wilms tumor images (MR) might have prognostic and histopathological diagnostic information with implications for the clinical assessment of response to chemotherapy. These data can then be used to optimize the Oncosimulator for predicting response to preoperative chemotherapy in Wilms tumor [9]. More research activities and clinical cases are required in order to confirm our preliminary hypothesis and we are expecting that in the nearest future we will strengthen our initial findings with additional evidence-based results.

ACKNOWLEDGMENT

Research supported by the European Commission under the project p-medicine (FP7-ICT-2009.5.3, Grant agreement number 270089).

REFERENCES

[1] Graf N, Tournade MF, de Kraker J: "The role of preoperative chemotherapy in the management of Wms' tumor. The SIOP studies." *Urol Clin North Am 27:443-454, 2000*

[2] Rosen G: "Neoadjuvant chemotherapy for osteogenic sarcoma: a model for the treatment of other highly malignant neoplasms." *Recent Results Cancer Res 103:148-157, 1986*

[3] DoctorEye platform web site (August 2012), http://biomodeling.ics.forth.gr.

[4] Skounakis E, Sakkalis V, Marias K A, Banitsas K, Graf N: "DoctorEye: A multifunctional open platform for fast annotation and visualization of tumors in medical images." *Conf Proc IEEE Eng Med Biol Soc. 2009:3759-3762, 2009.*

[5] The Contra Cancrum project web site (August 2012), http://www.contracancrum.eu

[6] Farmaki C, Marias K, Sakkalis V, Graf N. "Spatially adaptive active contours: a semi-automatic tumor segmentation framework." *Int J Comput Assist Radiol Surg. 2010 Jul;5(4):369-84.*

[7] Schenk JP, Graf N, Günther P, Ley S, Göppl M, Kulozik A, Rohrschneider WK, Tröger J. "Role of MRI in the management of patients with nephroblastoma." *Eur Radiol (2008) 18: 683–691, DOI 10.1007/s00330-007-0826-4*

[8] Rieden K, Weirich A, Tröger J, Gamroth AH, Raschke K, Ludwig R. "Accuracy of diagnostic imaging in nephroblastoma before preoperative chemotherapy." *Eur. Radiol. 3,115-122, 1993.*

[9] Stamatakos GS, Dionysiou DD, Graf NM, Sofra NA, Desmedt C, Hoppe A, Uzunoglu NK, Tsiknakis M: "The "Oncosimulator": a multilevel, clinically oriented simulation system of tumor growth and organism response to therapeutic schemes. Towards the clinical evaluation of in silico oncology." *Conf Proc IEEE Eng Med Biol Soc. 2007;2007:6629-32.*

Towards Patient Personalization of an Acute Lymphoblastic Leukemia Model during the Oral Administration of Prednisone in Children: Initiating the ALL Oncosimulator*

Eleftherios N. Ouzounoglou, *Student Member, IEEE,* Dimitra D. Dionysiou, Martin Stanulla and Georgios S. Stamatakos, *Member, IEEE*

***Abstract*— In the present study, methods aiming at supporting the personalization of an Acute Lymphoblastic Leukemia (ALL) Model (ALL Oncosimulator), already in development by the In Silico Oncology Group, National Technical University of Athens, are provided. Specifically, a population pharmacokinetic model for orally administered prednisone in children with ALL is developed, and the ability of classification algorithms to predict the prednisone response using gene expression data is studied.**

I. Introduction

Acute lymphoblastic leukemia (ALL) is one of the most common malignancies in childhood. Despite the significant advances made in the treatment of childhood ALL during the last decades, for approximately 25% of patients therapy still fails. Therefore, in conjunction with the forthcoming advances in the treatment of ALL, the development of multi-scale computational models is believed to contribute to further understanding of the disease and to the optimization of therapeutic strategies. Efforts on developing such a model, the ALL Oncosimulator, have already began by the In Silico Oncology Group (ISOG) of ICCS,NTUA, adopting similar modeling principals with already developed Oncosimulators [1].

Glucocorticoids, including Prednisone (PRED), belong to the main core of the treatment scheme followed in the ALL-BFM 2000 clinical trial and also in the majority of ALL- related chemotherapeutic schemes [2], since they are administered in almost every phase of treatment [3]. Apart from their therapeutic contribution, arising from the ability to induce apoptosis and cause cell cycle arrest in leukemic cells [2], response to PRED administration during the first 7 days of induction therapy (60 mg/m^2 of PRED orally given every 24 hours and one intrathecal dose of Methotrexate), is of crucial importance in order to categorize patients in high, intermediate or standard risk groups.

*This work was supported in part by the European Commission under the projects "p-medicine: From data sharing and integration via VPH models to personalized medicine" (FP7-ICT-2009.5.3), and "TUMOR: Transatlantic Tumour Model Repositories" (FP7- ICT-2009.5.4)

E.N. Ouzounoglou (elouzou@central.ntua.gr), D.D. Dionysiou (dimdio@esd.ece.ntua.gr) and Georgios S. Stamatakos (corresponding author) (phone: +30 210772 2287; fax: +30 2107733557; e-mail: gestam@central.ntua.gr) are with the Institute of Communication and Computer Systems, National Technical University of Athens, 9 Iroon Polytechniou, GR 157 80, Zografos, Greece

Martin Stanulla (e-mail: martin.stanulla@uk-sh.de), is with the Department of Pediatrics, University Medical Centre Schleswig-Holstein, Kiel, Germany.

Despite the significant role of PRED in the treatment of childhood ALL, the number of studies referring to the pharmacokinetic (PK) properties of PRED and its active metabolite Prednisolone (PREDNL) in children facing ALL is small, and even fewer studies deal with the modeling of their pharmacokinetics. Importantly, to our knowledge, a population PK modeling study for orally administrated PRED specifically in children with ALL has not been proposed yet. Such a model could allow the prediction of the pharmacokinetic behavior of PRED and PREDNL in patients enrolled in the ALL-BFM 2000 trial and the study of their pharmacodynamics in integration with the ISOG ALL Oncosimulator, in a personalized way. On the other hand, a population PK model of PREDNL in children with ALL has been proposed in [4], and various models for the PKs of orally administrated PRED are given in [5], [6]. So, a central target of the present study is to create a model capable of simulating the population PKs of PRED and PREDNL in childhood ALL patients based on existing related proposed models, due to the lack of PK data for orally administrated PRED in children with ALL.

Moreover, the highly significant knowledge of the Prednisone Response Group (PRG) of every patient, till today, could only be inferred after the first 7 days of PRED treatment by measuring the number of peripheral blood leukemic blasts persisting after the induction therapy phase. A priori knowledge of the PRED response category would enable clinicians to promptly optimize the treatment in a patient-specific manner. This knowledge would be also considerably informative for the procedures of personalized initialization and adaptation of the ALL Oncosimulator's input parameters, since it has been shown that leukemic cells exhibit differential inter-patient characteristics as far as PRG is concerned (e.g. both the induction of glucocorticoid induced apoptosis and cell cycle arrest are more evident in good response group [7], [8]). Up to now, significant efforts have been made in order for the genes showing differential expression before and after the initial treatment of childhood ALL to be identified between poor and good responders [9–13]. Moreover, the prediction of the outcome of initial phases of treatment has been attempted by using the expression data of specific genes [9], [11]. However, to our knowledge, the prediction of the PRG using gene expression data has not been attempted yet, especially using the available (within the context of the p-medicine project (FP7-ICT-2009.5.3)) set of genes (86 in number), which

derives from the acquired knowledge for the differentially expressed genes between the two groups described in [12].

The present study consists of two distinct sections: one for the derivation of a PK model for orally administered PRED in children with ALL and one for the study of the capability of various classification algorithms to predict the PRG of patients based on gene expression data. Although these two tasks have been treated autonomously, they share the same target of contributing to the personalization of the ALL Oncosimulator.

II. Oral Prednisone Population Pharmacokinetics in Children with ALL

A. Existing Prednisone Related Pharmacokinetic Modeling Studies

As previously described, a number of studies have dealt with the PK modeling of PRED and/or its active metabolite PREDNL. However, for the creation of the newly proposed PK model of oral PRED only some of them were chosen to be combined.

In [4] the modeling of population PKs of orally or intravenously administered PREDNL in children with ALL has been studied. Briefly, a 2-compartmental PK model for oral PREDNL has been proposed, where the drug is absorbed by a first order process in the Central Compartment (of volume V_c). After the absorption, the drug can be reversibly transferred to the Peripheral Compartment (of volume V_p) (Inter-compartmental clearance rate: CL_d) or eliminated from the Central Compartment (Clearance rate: CL). Together with the pharmacokinetic model, a regression model able to predict the values of PK parameters (CL, V_c, V_p), fitted to PK data of children with ALL has been given, having as input the Body Surface Area (BSA) and the Weight (WT) of the patient. This regression model is of crucial importance since it permits the estimation of the basic pharmacokinetic parameters of the active metabolite of PRED, PREDNL, in a personalized way.

In [5] a model for the prediction of PRED and/or PREDNL levels in human plasma after either oral or intravenous administration of PRED or PREDNL was proposed. In this model either PRED or PREDNL are firstly absorbed in a depot compartment in a percentage of the given dose, determined by the bioavailability of the drug. A fraction of the drug is absorbed by a first order process in its parent form (PRED or PREDNL) and the remainder in the metabolized form (PREDNL or PRED) due to the first pass conversion of the drug. Two additional compartments concern the volume of distribution of the two forms of the drug which could further be reversibly metabolized. Both forms of the drug could be eliminated from the related compartment.

B. Development of a Pharmacokinetic model for orally administered Prednisone in children with ALL

As already stated, the acquirement of a PK model capable for simulating the pharmacokinetics of PRED in children with ALL is of crucial importance for the accurate simulation of the related therapy by the ALL Oncosimulator. However, such a model, to our knowledge, does not exist, and the lack of data for the levels of PRED and PREDNL in blood plasma of children with ALL after oral administration of PRED does not allow the development of a new population PK model dealing with the details of PRED PKs such as PRED and PREDNL levels, their reversible metabolism, bioavailability and other significant PK parameters. So, an already derived population PK model [4], has been reutilized in order to simulate the levels of the active metabolite of PRED, PREDNL in a personalized way. However some specific extensions of this model had to be made in order to be capable for simulating the absorption of orally administrated PRED as well. The modeling principles for these changes, were acquired from the model presented in [5].

1) The structure of the extended proposed model

The structure of the proposed oral PRED PK model, combining parts of previously presented models ([4], [5]) is shown in Figure 1. Orally administered PRED is firstly absorbed in a depot compartment in a percentage of the given dose, determined by the bioavailability (F) of the drug, as in [5]. A fraction of the drug (f_{PREDNL}) is absorbed in its metabolized form (PREDNL), due to its instant first-pass conversion, by the Central Compartment, which is identical to the Central Compartment in [4], by a first order process (ka). Similar to [4], the absorbed drug could be reversibly transferred (inter-compartmental clearance rate: CLd) to Peripheral Compartment (which is also identical to the Peripheral Compartment in [4]) or eliminated from Central Compartment (Clearance rate: CL). The assumption made here is that the major part of conversion from PRED to PREDNL is already done by the first pass conversion. So, the following reversible metabolism of PRED to PREDNL and the levels of PRED in the two compartments (Central and Peripheral) are not modeled, a choice also made in [4] for the administration of PREDNL. Moreover, the same

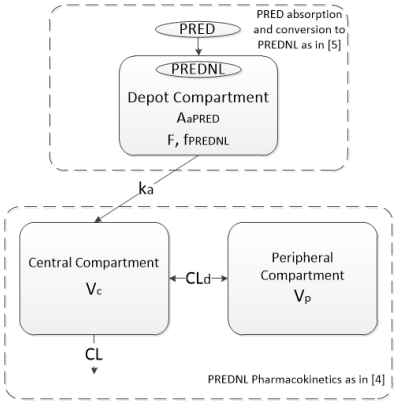


Figure 1. The structure of the proposed pharmacokinetic model for orally administered Prednisone.

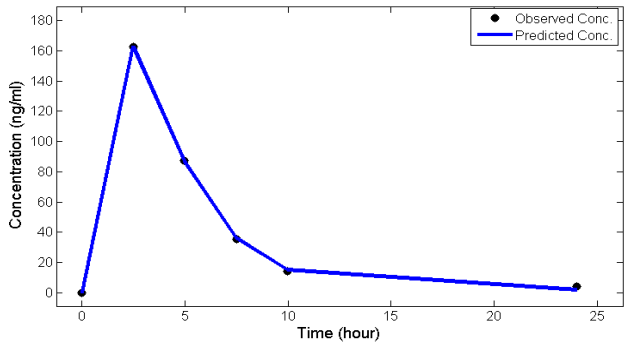


Figure 2. The results of the proposed pharmacokinetic model simulation after the parameter estimation procedure using the data given in [14]. The solid black circles refer to the measurements and the blue line refers to the predicted concentrations of the drug in the Central Compartment.

assumption was made in [6] for the case of orally administered PRED.

2) Evaluation of the model using Pharmacokinetic data

In order to ensure that the proposed model, as far as its structure is concerned, is capable for simulating the PKs of PREDNL after oral administration of PRED, PK data from children treated with oral PRED have been used for parameter fitting. Unfortunately, as already mentioned, such data could not be found in literature for children with ALL. However, data from children with nephrotic syndrome, treated with oral PRED are given in [14]. The measurements taken in remission phase were used, although it is shown by the same study that the prednisolone levels during the active phase did not differ significantly from those observed during remission. In order for the figure depicting the mean concentration of free PREDNL after oral administration of 55mg of PRED measured in [14] to be digitized and quantified, the GetData Graph Digitizer 2.25 software was used and the acquired data are given in Figure 2.

The fitting of the model was done by using the Parameter Fit functionality of MATLAB 2012a (The MathWorks Inc., 2012, Natick, Massachusetts) SimBiology Toolbox (individual fit using NLINFIT function). In order to acquire initial values for the pharmacokinetic parameters, the regression model of [4] was used. It should be noted that the initial estimate for V_c was slightly changed from the value that the regression model dictates in order for a first adequate fit to the data to be achieved by manual fine tuning. A higher value in V_c is expected due to lower (even in remission phase) serum albumin concentration found in patients in [14] (29 g/l) in contrast with the patients in [4] (36 g/l). After the parameter estimation procedure, the PK model was able to accurately simulate the dynamics of PREDNL levels in the Central Compartment as shown in Figure 2.

III. Prednisone Response Prediction

A. Available data, preprocess analysis, classification algorithms and gene selection methods

In the context of p-medicine project (FP7-ICT-2009.5.3), data for a number of 665 patients, enrolled in the ALL-BFM 2000 clinical trial, has been made available. Among others, the expression levels for a set of 86 genes that are recurrently identified as being differentially expressed between Good and Poor Prednisone responders [12] were given. Although considerable efforts have been made in order to identify the genes that are differentially expressed between these groups [10], [12], [13], the given set has not been studied for the statistical significance of differential expression and for the ability to derive a classifier able to predict the response group. Therefore, the main target of the present task was to study the predictability based on the given set of genes and the statistical significance of their potential differential expression as well, in order to ensure that the findings do not apply only to the observed dataset. As far as preprocess analysis from patient datasets are concerned, the example of similar previous studies [9], [11] has been followed. More specifically, only patients facing precursor B-ALL, negative for chromosomal rearrangements TEL-AML1, BCR-ABL an MLL-AF4 were included, and patients showing DNA index different from 1.0 were excluded. This first separation of data resulted in 259 prednisone good responders (PGR) and 26 prednisone poor responders (PPR). However, in order to avoid statistical bias in the first class, a random choice of 26 patients from the PGR category has been made trying to match every patient from the poor response category to a patient with same characteristics for age, sex and white blood cells (WBC) count from the good response category, a step proposed and followed in [12]. The final sets of data were introduced into MATLAB. Any missing expression value was filled by using the *knnimpute* function which replaces the missing values with a weighted mean of the k nearest-neighbor (the value of k was chosen to be 10).

The classification algorithms used were the Discriminant Analysis (DA) algorithm implemented in MATLAB by the

TABLE I. Prediction Accuracy Results of the Derived Classifiers using the "Leave One Out" Method on the Available Data

Classifier (Type)	DA(lin.)	DA(dlin.)	DA(quad.)	DA(dquad.)	SVM(lin.)	SVM(quad.)	SVM(poly.)	SVM(RBF)	SVM(MLP)
Gene Selection Method	**Prediction Accuracy,**Sensitivity,Specificity **(Leave One Out Method)**								
Entire Gene List (86 genes)	not applicable	**0.61**,0.60,0.62	not applicable	**0.69**,0.66,0.72	**0.88**,0.85,0.91	**0.53**,0.53,0.53	**0.71**,0.67,0.76	**0**,0,0	**0.55**,0.57,0.54
T-test & FDR (3 genes)	**0.84**,0.80,0.90	**0.84**,0.80,0.90	**0.79**,0.88,0.69	**0.78**,0.74,0.85	**0.80**,0.76,0.86	**0.71**,0.68,0.73	**0.71**,0.68,0.73	**0.76**,0.75,0.79	**0.75**,0.72,0.78
Rank Products (13 genes)	**0.76**,0.81,0.73	**0.65**,0.65,0.65	**0.73**,0.71,0.75	**0.65**,0.62,0.70	**0.71**,0.72,0.70	**0.59**,0.61,0.58	**0.61**,0.69,0.58	**0.57**,0.57,0.57	**0.61**,0.60,0.62
Gene Selector (9 genes)	**0.80**,0.78,0.83	**0.86**,0.80,0.95	**0.71**,0.78,0.66	**0.73**,0.83,0.67	**0.86**,0.80,0.95	**0.63**,0.64,0.62	**0.69**,0.67,0.70	**0.67**,0.68,0.66	**0.75**,0.74, 0.76

classify function (using linear, diaglinear, quadratic and diagquadratic options) and the Support Vector Machine algorithm implemented in MATLAB by the *svmtrain* and *svmclassify* functions (using linear, quadratic, polynomial, rbf and mlp options for kernel). In order to evaluate the performance of the classifiers, the Leave One Out method was used. In order for the classifiers to be accepted as a tool able to predict the category of a given patient, we had to ensure that their discriminative power is not limited only in the dataset used for their training, by confirming that the classification features used, here the gene expression values, show statistically significant differences between the groups. Therefore, before training the classification structures, various statistical tests have been made which were: the two sample unpaired t-test and the estimation of false discovery rate (FDR) for multiple hypothesis testing (using mattest and mafdr functions in MATLAB), the Rankproducts method (using Bioconductor package 'RankProd' in R (R Foundation for Statistical Computing, Vienna, Austria)) and the GeneSelector method (using Bioconductor package 'GeneSelector' in R), combining three well established rankings methods together with the t-test for unequal variances (Welch's T-statistic in Gene Selector package), the parametric empirical Bayes (limma package), the SAM statistic (samr package) and the Permutation test (multtest package).

B. *Classification Training and Testing prior and after the Gene Selection*

The results of prediction accuracy after training the classifiers with the different sets of genes as classification features are shown in Table I. The results of the Leave One Out method are presented because it is believed to be the most significant criterion, since in a real life scenario, a single new patient should be classified based on the knowledge acquired by all the previously encountered patients. As it can be seen for the case of using the Gene Selector feature selection and the SVM classification algorithm with linear kernel, a prediction accuracy of 86.54% with particularly good sensitivity and specificity statistics is achieved, which does not deviate significantly from the performance of the same algorithm prior to the gene selection step. This, as already stated, strengthens the applicability of the classification scheme to the population from which the data samples were taken.

IV. Conclusion

In the present study, both the development of a pharmacokinetic model for orally administered prednisone in children with ALL has been achieved and the ability for a classification scheme able to predict the Prednisone Response Category of patients has been investigated. The results showed that the proposed pharmacokinetic model is capable for simulating the time evolution of the concentration levels of free prednisolone after oral administration of prednisone in children. Moreover, the prediction accuracy of the derived classifiers is reasonably good so as to support the use of such a tool for the Prednisone Response Category prediction. Therefore, the proposed tools presented in this study are believed to be able to support the procedure of parameterization of the ALL Oncosimulator in a personalized way.

References

[1] G. Stamatakos, "*In Silico* Oncology Part I: Clinically Oriented Cancer Multilevel Modeling Based on Discrete Event Simulation," in *Multiscale Cancer Modeling*, T. Deisboeck and G. Stamatakos, Eds. Boca Raton, Florida, USA: CRC Press, 2010, pp. 407–436.

[2] H. Inaba and C.-H. Pui, "Glucocorticoid use in acute lymphoblastic leukaemia," *The Lancet Oncology*, vol. 11, no. 11, pp. 1096–1106, Nov. 2010.

[3] G. Basso, M. Veltroni, M. G. Valsecchi, M. N. Dworzak, R. Ratei, D. Silvestri, A. Benetello, B. Buldini, O. Maglia, G. Masera, V. Conter, M. Arico, A. Biondi, and G. Gaipa, "Risk of relapse of childhood acute lymphoblastic leukemia is predicted by flow cytometric measurement of residual disease on day 15 bone marrow," *Journal of Clinical Oncology*, vol. 27, no. 31, pp. 5168–5174, Nov. 2009.

[4] K. B. Petersen, W. J. Jusko, M. Rasmussen, and K. Schmiegelow, "Population pharmacokinetics of prednisolone in children with acute lymphoblastic leukemia," *Cancer Chemotherapy and Pharmacology*, vol. 51, no. 6, pp. 465–473, Jun. 2003.

[5] J. Xu, J. Winkler, and H. Derendorf, "A pharmacokinetic/pharmacodynamic approach to predict total prednisolone concentrations in human plasma," *Journal of Pharmacokinetics and Pharmacodynamics*, vol. 34, no. 3, pp. 355–372, Jun. 2007.

[6] M. H. Magee, R. A. Blum, C. D. Lates, and W. J. Jusko, "Pharmacokinetic/pharmacodynamic model for prednisolone inhibition of whole blood lymphocyte proliferation," *British Journal of Clinical Pharmacology*, vol. 53, no. 5, pp. 474–484, May 2002.

[7] F. Casale, R. Addeo, V. D'Angelo, P. Indolfi, V. Poggi, C. Morgera, S. Crisci, and M. T. Di Tullio, "Determination of the in vivo effects of prednisone on Bcl-2 family protein expression in childhood acute lymphoblastic leukemia.," *International Journal of Oncology*, vol. 22, no. 1, pp. 123–128, 2003.

[8] R. Addeo, F. Casale, M. Caraglia, V. D'Angelo, S. Crisci, A. Abbruzzese, M. T. Di Tullio, and P. Indolfi, "Glucocorticoids induce G1 arrest of lymphoblastic cells through retinoblastoma protein Rb1 dephosphorylation in childhood acute lymphoblastic leukemia in vivo," *Cancer Biology & Therapy*, vol. 3, no. 5, pp. 470–476, May 2004.

[9] D. Bhojwani, H. Kang, R. X. Menezes, W. Yang, H. Sather, N. P. Moskowitz, D.-J. Min, J. W. Potter, R. Harvey, S. P. Hunger, N. Seibel, E. A. Raetz, R. Pieters, M. A. Horstmann, M. V. Relling, M. L. den Boer, C. L. Willman, and W. L. Carroll, "Gene expression signatures predictive of early response and outcome in high-risk childhood acute lymphoblastic leukemia: A Children's Oncology Group Study [corrected].," *Journal of Clinical Oncology*, vol. 26, no. 27, pp. 4376–84, Sep. 2008.

[10] P. Rhein, S. Scheid, R. Ratei, C. Hagemeier, K. Seeger, R. Kirschner-Schwabe, A. Moericke, M. Schrappe, R. Spang, W.-D. Ludwig, and L. Karawajew, "Gene expression shift towards normal B cells, decreased proliferative capacity and distinct surface receptors characterize leukemic blasts persisting during induction therapy in childhood acute lymphoblastic leukemia," *Leukemia: Official Journal of the Leukemia Society of America, Leukemia Research Fund, U.K*, vol. 21, no. 5, pp. 897–905, May 2007.

[11] G. Cario, M. Stanulla, B. M. Fine, O. Teuffel, N. V. Neuhoff, A. Schrauder, T. Flohr, B. W. Schäfer, C. R. Bartram, K. Welte, B. Schlegelberger, and M. Schrappe, "Distinct gene expression profiles determine molecular treatment response in childhood acute lymphoblastic leukemia," *Blood*, vol. 105, no. 2, pp. 821–826, Jan. 2005.

[12] G. Cario, A. Fetz, C. Bretscher, A. Möricke, A. Schrauder, M. Stanulla, and M. Schrappe, "Initial leukemic gene expression profiles of patients with poor in vivo prednisone response are similar to those of blasts persisting under prednisone treatment in childhood

acute lymphoblastic leukemia," *Annals of Hematology*, vol. 87, no. 9, pp. 709–716, Sep. 2008.

[13] W. J. E. Tissing, M. L. den Boer, J. P. P. Meijerink, R. X. Menezes, S. Swagemakers, P. J. van der Spek, S. E. Sallan, S. A. Armstrong, and R. Pieters, "Genomewide identification of prednisolone-responsive genes in acute lymphoblastic leukemia cells," *Blood*, vol. 109, no. 9, pp. 3929–3935, May 2007.

[14] G. Gatti, E. Perucca, G. M. Frigo, L. D. Notarangelo, L. Barberis, and A. Martini, "Pharmacokinetics of prednisone and its metabolite prednisolone in children with nephrotic syndrome during the active phase and in remission," *British Journal of Clinical Pharmacology*, vol. 17, no. 4, pp. 423–431, Apr. 1984.

Quantification of Cytotoxic Chemotherapy Effects Within the Linear-quadratic Model of Radiotherapy.

Roger G. Dale

Abstract— **The use of chemoradiotherapy (CRT) is widespread but quantitative evaluation of the benefit accruing from concomitant use of cytotoxic drugs is difficult since radiation and chemotherapy "doses" are currently expressed in different and incompatible ways. As the biological impact of any radiotherapy schedule may be expressed in terms of a biologically effective dose (BED) it is shown that chemotherapy effects can be similarly quantified, allowing drug contributions to be expressed in terms of a radiation-equivalent quantity which is familiar to radiation oncologists. This paper discusses various ways in which chemotherapy can enhance radiation effect and derives the associated BED equations. The availability of such equations can potentially help inform clinical judgement in relation to the true value of CRT treatments and also allows derivation of alternative, radiation-only, schedules.**

I. INTRODUCTION

The combined use of concomitant cytotoxic chemotherapy and radiotherapy (chemoradiotherapy, CRT), is widespread. Chemotherapy agents can provide useful gains (typically 10 - 20%) in tumour response, the improvement coming mostly from increased local tumour control probability (TCP) control rather than a decrease in distant metastases. Although there exists an extensive literature relating to the modelling of cases where cytotoxic drugs are used in isolation (e.g. [1,2]). There may be several mechanisms by which drugs can enhance radiation cell kill and at present these are not well-understood. Additionally, radiotherapy and chemotherapy "doses" are currently specified in entirely different and incompatible ways, thus providing a major additional complication when attempting to evaluate CRT.

CRT is favoured if the resultant loco-regional control is better than that achieved by either chemotherapy or radiotherapy alone and *for comparable normal tissue toxicity*. However, one of the difficulties in evaluating CRT lies with the fact that the addition of chemotherapy may bring new patterns of toxicity not associated with radiation alone. In this regard the design of clinical trials has not always been especially good since ideally a minimum three-arm strategy (radiotherapy alone at dose D1, radiotherapy at dose D1 + drug; radiotherapy alone at higher dose D2) is ideally required.

There is thus a need to attempt to evaluate CRT in terms of parameters which link directly to the large volume of clinical experience which has accrued from the use of radiotherapy alone. One way of achieving this is by application of the linear-quadratic (LQ) model of radiation effect, which is widely used for the radiobiological quantification of radiotherapy [3-5]. Originally developed for the analysis of conventional fractionated treatments, it has since been extended to cover many other patterns of radiation delivery, e.g. continuous low dose-rate therapy, permanent implants, targeted radionuclide therapy, etc [6,7]. On account of the growing worldwide interest in particle therapy (i.e. treatments involving protons, neutrons, carbon ions, etc), the LQ model has also been adapted to be used for quantifying high-LET (Linear Energy Transfer) radiations [8,9].

R.G.Dale is with the Faculty of Medicine, Imperial College, London, UK. (Tel: 44+(0)2085671317; email: r.dale@imperial.ac.uk).

There are sound reasons for wishing to express chemotherapy effects in terms of a radiation-equivalent. Modern radiotherapy enjoys high standards of dosimetric accuracy and this gives added credence to the results of clinical trials. Without high dosimetric standards the intercomparisons between different trials and schedules would be mostly unreliable and radiobiological modelling would be of little value. Present chemotherapy dose prescription practices do not remotely approach the high standards routinely expected from radiotherapy and it thus makes sense to attempt to quantify drug therapy results in radiation-equivalent measures which are familiar to clinical oncologists.

In this article some possible mechanisms of drug-enhancement are considered in terms of existing LQ methodology.

II. BIOLOGICALLY EFFECTIVE DOSE (BED) AND ITS RELATIONSHIP TO LOG CELL KILL.

For a fractionated radiotherapy treatment involving N well-spaced treatment fractions, each delivering a fixed dose (d), the cellular surviving fraction (S) is given by:

$$S = \exp\left(-N\alpha d - N\beta d^2\right) \tag{1}$$

where α (in units of Gy^{-1}) and β (units of Gy^{-2}) are the respective linear and quadratic radiosensitivity coefficients in the LQ model. For clinical application it is usually more efficient to express treatment in terms of biologically effective dose (BED), defined as follows [Barendsen; Fowler]:

$$BED = -\frac{\ln S}{\alpha} = Nd\left[1 + \frac{d}{\left(\frac{\alpha}{\beta}\right)}\right] \tag{2}$$

BED is the physical dose required to produce the observed biological response if the treatment were to be given in an infinite number of very small fractions. It will be noted that BED is dependent on the ratio α/β (units of Gy),

rather than on the values of the individual radiosensitivity parameters. As α/β values (sometimes referred to as fractionation factors) are tissue-specific, then so also is BED. A typically-used generic value of α/β for relatively fast-growing tumours (eg squamous cell cancers, SCCs) is 10Gy. For late -responding normal tissues a generic α/β value of 3Gy is conventionally applied a,though, in both cases, it should be noted that there are important exceptions to the general rule. Calculated BEDs are usually written with a subscript (3, 10, etc) which denotes the specific α/β value assumed in the derivation.

For tumours which exhibit significant repopulation during treatment a subtractive factor may be added to the right hand side of (2) to reflect the reduced BED which results [4,5,10]. The form of the subtractive repopulation factor is:

$$-k\left(T-T_{delay}\right) \tag{3}$$

where T is the overall treatment time, T_{delay} is the time lag from initiation of treatment before which the repopulation begins and k is the BED equivalent of repopulation (in units of $Gyday^{-1}$). Since this article confines itself to general principles the repopulation effect will not be considered in depth here.

For what follows it is useful to gauge the nature of the relationship between BED and log cell kill. Re-writing the BED definition in (2) yields:

$$S = e^{-\alpha BED} \tag{4}$$

and, if S is expressed in the conventional logarithmic fashion as a number (X) of log cell kills (i.e. as 10^{-X}), then:

$$10^{-X} = e^{-\alpha BED} \tag{5}$$

leading to:

$$BED = \frac{2.303X}{\alpha} \tag{6}$$

Table 1 shows the relationship between log cell kill and BED for tumours posessing average radiosensitivity (i.e. $\alpha = 0.35Gy^{-1}$) and a generic α/β value of 10Gy. The table additionally tabulates the equivalent number of 2Gy fractions required to achieve the given BEDs, as calculated via (2). Alternative tabular values for other assumed parameter values may be easily derived.

Depending on circumstances, conventional radiation-only schedules typically deliver BEDs which would be expected to produce cell kills in the wide range 10^{-8} to 10^{-12}. Since the additional cell kill due to cytotoxic chemotherapy can, in principle, also be expressed in a similar manner to that described in Table I, this opens the way to expressing the chemotherapy component of a CRT treatment in terms of a radiation BED-equivalent, or in terms of an equivalent number of 2Gy fractions [11].

Chemotherapy alone rarely achieves a cell kill of better than $< 10^{-6}$, with a number of studies suggesting that the true cell kill is in many cases very much less than this figure, i.e. around 10^{-2}-10^{-3}. This may seem quite modest but, in terms of improved local control, could nonetheless be very beneficial. Such a degree of cell kill may be ineffective for dealing with microscopic metastases, however, thus challenging one of the main claims often made for adjuvant chemotherapy [12].

TABLE I. RELATIONSHIP BETWEEN LOG CELL KILL, BED AND EQUIVALENT NUMBER OF 2-GY FRACTIONS FOR TUMOURS OF AVERAGE RADIOSENSITIVITY (I.E. $\alpha = 0.35Gy^{-1}$) AND WITH $\alpha/\beta = 10Gy$.

Cell kill	BED	Equivalent number of 2 Gy fractions [from (2)]
10^{-1}	6.6	2.8
10^{-2}	13.2	5.5
10^{-3}	19.7	8.2
10^{-4}	39.5	16.5
10^{-10}	65.8	27.4

A further useful feature of BED is that it is an additive quantity, i.e. the BED associated with two different treatment types may be added to determine the resultant BED. Thus, irrespective of any assumed mechanisms, drug-induced cell kill can, in principle, always be expressed in terms of an equivalent radiation-only BED through the analysis of clinical outcomes. For example, suppose local control was found to be identical in the following two cases:

33 × 2Gy (radiation alone)

and:

26 × 2.25Gy radiation + cytotoxic chemotherapy.

Then, on account of BED additivity:

$$BED_{CRT} = BED_{RT} + BED_{chemo} \tag{7}$$

and, since the CRT and radiation-only schedules are iso-effective, the drug-induced BED (BED_{chemo}) is simply the difference between the two radiation BEDs, i.e:

$$BED_{chemo} = 33\times 2\times\left[1+\frac{2}{10}\right] - 26\times 2.25\times\left[1+\frac{2.25}{10}\right]$$

i.e. $BED_{chemo} = 7.5Gy_{10}$, corresponding to what would be expected from around three 2Gy radiation fractions.

III. SOME POSSIBLE MODES OF CHEMOTHERAPY ENHANCEMENT OF RADIATION EFFECTS.

Improved local control arising from the chemotherapy component of CRT may be a result of any one, or combination of, the following:

i) Independent cell kill.

This implies that the chemotherapy sterilises a sub-population of cells which may be resistant to radiation-induced cell kill. In this case the resultant tumour BED

(BED_{CRT}) is simply given by (7). Within this general heading other possibilities, e.g. cell-cycle synchronisation, enhanced apoptosis or re-oxygenation may all intrinsically contribute to the general phenomenon of independent cell kill.

ii) Direct dose sensitisation.

This would imply a drug-radiation interaction which effectively increments each fractional radiation dose. If the degree of enhancement is s, then BED_{CRT} is derived from (2) as:

$$BED_{CRT} = Nsd\left[1+\frac{sd}{\left(\frac{\alpha}{\beta}\right)}\right] \tag{8}$$

In this case the chemotherapy sensitisation acts in a non-linear way since it increments both the total dose (Nd) and the fraction dose (d).

Since the expected radiation-only BED (without drug enhancement) is given by (2), the extra drug-induced BED inherent in (8) is the difference between the two expressions, i.e:

$$BED_{chemo} = Nsd\left[1+\frac{sd}{\left(\frac{\alpha}{\beta}\right)}\right] - Nd\left[1+\frac{d}{\left(\frac{\alpha}{\beta}\right)}\right] \tag{9}$$

i.e.:

$$BED_{chemo} = Nd(s-1)\left[1+\frac{d(s+1)}{\left(\frac{\alpha}{\beta}\right)}\right] \tag{10}$$

[It should be noted that, even if a subtractive repopulation factor (3), is included in (2) and (8), it would cancel out in arriving at (10)]. Thus, for a 35 × 2Gy radiation schedule, a modest dose sensitisation factor of 5% (ie s = 1.05) would produce a BED_{chemo} value of ~$4Gy_{10}$. From Table I this is seen to be compatible with a drug-induced cell kill of less than 10^{-1}, which could alternatively be achieved with two extra 2Gy fractions of radiation.

iii) Sensitisation of the radiation effect by increasing the α and/or β radiosensitivity parameters.

This possibility differs from pure dose sensitisation and is exactly analogous to that associated with the RBE enhancements which result from the use of radiations possessing a high linear energy transfer (LET). When such radiations are deployed the conventional (i.e. low-LET) α-radiosensitivity parameter is increased from α_L to α_H by a factor RBE_{max}, where:

$$RBE_{\max} = \frac{\alpha_H}{\alpha_L} \tag{11}$$

The subscripts in (1) relate to the low-and high-LET cases. RBE_{max} is the maximum RBE that would be expected with the high-LET radiation when dose fractions are very small. RBE itself is a variable which maximised at low fraction dose and falls asymptotically to a lower limiting value (RBE_{min}) at high fraction dose. RBE_{min} is often assumed to be unity, but there is evidence that it may be non-unity in some cases [9,13]. In such cases the lower limiting value of RBE is related to a changed β-radiosensitivity parameter and is given by [9]:

$$RBE_{\min} = \sqrt{\frac{\beta_H}{\beta_L}} \tag{12}$$

It has been shown [9] that incorporation of (10) and (11) leads to a modified version of (2) which is applicable to high-LET therapy:

$$BED_H = N_H d_H\left[RBE_{\max} + \frac{RBE_{\min}^2 d_H}{\left(\frac{\alpha}{\beta}\right)_L}\right] \tag{13}$$

Returning to the chemotherapy case: if a cytotoxic drug increases the low-LET α-value by a factor A, then A influences the LQ formulation in the same manner as does RBE_{max}, as discussed above. Similarly, if the drug increases the low-LET √β-parameter by a factor B, then B influences the formulation in an analogous manner to RBE_{min}. In the case where A and B enhancements are both operative then:

$$BED_{CRT} = Nd\left[A+\frac{B^2 d}{\left(\frac{\alpha}{\beta}\right)}\right] \tag{14}$$

The similarity of form between (13) and (14) will be noted. If, for example, B=0, then the chemotherapy BED enhancement associated with (14) is [following the same procedure used to derive (10)]:

$$BED_{chemo} = Nd\left[(A-1)+\frac{d\left(B^2-1\right)}{\left(\frac{\alpha}{\beta}\right)}\right] \quad (15)$$

Comparing (10) and (15) it will be noted how pure dose sensitisation and radiosentivity-sensitisation lead to BED expressions with differing forms, the two equations being identical only in the special case where s=A=B. If only one of the sensitisation routes is operative then (14) and (15) are modified simply by setting either A or B equal to unity, as appropriate.

In the case where CRT involves a high-LET radiation component, and if it is assumed that all the modifiers (A, B, RBE_{max} and RBE_{min}) are operative then (14) becomes:

$$BED_{CRT} = Nd\left[ARBE_{\max}+\frac{B^2 RBE_{\min}^2 d}{\left(\frac{\alpha}{\beta}\right)}\right] \quad (16)$$

As with (13) and (14) the individual sensitising and RBE factors may be set to unity in cases where that is appropriate.

iv) Reduction or inhibition of tumour repopulation during the inter-fraction intervals.

This phenomenon could be a result of a reduction in the value of k, the BED-equivalent repopulation factor, or an increase in the lag time (T_{delay}) before which fact repopulation becomes significant, in which case then one or both of these parameters would take altered values in (10). If chemotherapy is able to completely inhibit repopulation then k=0 and (1) would not be required.

IV. Estimation of Chemotherapy Enhancement from Knowledge of Changed TCP.

Observed TCPs may also be used in the direct evaluation of a therapy radiation-equivalent. If any treatment schedule leads to an average surviving cell number of n then, from Poisson statistics:

$$TCP = e^{-n} \quad (17)$$

If the initial clonogen number is C then, for a radiation-only schedule:

$$n = Ce^{-\alpha BED_{RT}} \quad (18)$$

and:

$$TCP_{RT} = e^{-Ce^{-\alpha BED_{RT}}} \quad (19)$$

Similarly, for a CRT schedule, and bearing in mind the additivity of the BEDs of the respective contributions (7):

$$TCP_{CRT} = e^{-Ce^{-\alpha\left(BED_{RT}+BED_{chemo}\right)}} \quad (20)$$

Combining (19) and (20) leads to:

$$BED_{chemo} = \frac{1}{\alpha}\ln\left[\frac{\ln\left(TCP_{RT}\right)}{\ln\left(TCP_{CRT}\right)}\right] \quad (21)$$

a result which is independent of C and hence of initial tumour size. Thus, assuming an α value of $0.35Gy^{-1}$, an observed increase in TCP from 0.2 to 0.5 would be consistent with a BED_{chemo} value of 2.41Gy, which is equivalent to the addition of one 2Gy fraction. This result again indicates that even modest drug contributions can translate to potentially very valuable improvements in treatment outcome. Application of (8) or (15) allows subsequent estimation of the specific chemotherapy enhancement factors from such results.

V. Discussion

LQ methodology is well-established in radiation oncology and it is demonstrated here that the same principles can be used to glean useful quantitative estimations of the value of the concurrent addition of cytotoxic chemotherapy to radiotherapy. Providing radiation-based benchmarks to such deliberations is valuable, particularly as it appears that drug contributions, although often contributing to improved local control, may nonetheless be providing only a very modest additional degree of cell kill. The expression of a drug contribution in terms of a radiation-equivalent allows judgement to be made over whether the same enhancement in local control would be better achieved by a modest increase in radiation dose. If this were feasible via one of the forms of focal (e.g. Intensity Modulated or particle-) radiotherapy then any confounding issues associated with drug-induced changes to the pattern of normal tissue toxicity might be avoided.

The formulations developed here are not restricted to drug applications and could, in principle, be applied to any other form of therapy which can be used in combination with radiation. Some steps have already been taken in using this approach in the assessment of hyperthermia [14].

VI. Conclusion

The expression of chemotherapy effects in terms of a radiation BED-equivalent could allow easier correlation between clinical experience with CRT and the wider range of radiation-only treatments. Although at present it is not possible to ascertain exactly which mechanism(s) might contribute to the chemotherapy cell kill there is a

straightforward relationship between BED and cell surviving fraction. The methodology discussed here could therefore have interesting application in analysing the results of clinical trials. Additionally, the incorporation of the methodology to *in-silico* modelling (albeit in slightly modified form) could allow a more detailed examination of the various modes of drug cell kill which might operate in CRT [15].

ACKNOWLEDGMENTS

THE AUTHOR WISHES TO THANK GEORGE PLATANIOTIS, ALEJANDRO CÁRABE-FERNÁNDEZ, BLEDDYN JONES, JACK FOWLER AND ALAN NAHUM FOR STIMULATING DISCUSSIONS.

REFERENCES

[1] M. Joerger, A.D. Huitema, S. Krähenbühl, J.H. Schellens, T. Cerny, T. Reni, M. Zucca, F. Cavalli and A.J. Ferreri, "Methotrexate are-under-the-curve is an important outcome predictor in patients with primary CNS lymphoma: A pharmacokinetic-pharmacodynamic analysis from the IELSG no 20 trial," *Br J Cancer* vol. 102, no. 4, pp673-677, February, 2010.

[2] L.G. Marcu and W.M. Harriss-Phillips, "In-silico modelling of tretament-induced tumour cell kill: developments and advances," *Comput Math Methods Med*, Epub 960256, July, 2012.

[3] G. W. Barendsen, "Dose fractionation, dose-rate and iso-effect relationships for normal tissue response," *Int J Radiat Oncol Biol Phys* 8, pp1981-1997, 1982.

[4] J.F Fowler, "The linear-quadratic formula and progress in fractionated radiotherapy," *Br J Radiol* vol.62, pp679-694, August, 1989.

[5] J. F. Fowler, "21 years of biologically effective dose," *Brit J Radiol* vol. 83, pp554-568, July, 2010.

[6] R. G. Dale, "The application of the linear-quadratic dose-effect equation to fractionated and protracted radiotherapy," *Brit J Radiol* vol.58, pp515-528, June, 1985.

[7] B. Jones, R.G. Dale, C. Deehan, K.I. Hopkins and D.A Morgan, "The role of biologically effective dose (BED) in clinical oncology," *Clin Oncol* 13, pp71-81, 2001.

[8] R.G. Dale and B. Jones. "The assessment of RBE effects using the concept of biologically effective dose," *Int J Radiat Oncol Biol Phys* vol.43, no. 3, pp639-645, 1999.

[9] A Cárabe-Fernández, R.G. Dale and B.Jones, "The incorporation of the concept of minimum RBE (RBEmin) into the linear-quadratic model and the potential for improved radiobiological analysis of high-LET treatments," *Int J Radiat Oncol Biol Phys* vol.83, no. 1, pp27-39, January, 2009.

[10] E.L. Travis and S.L. Tucker, "Isoeffect models and fractionated radiotherapy," *Int J Radiat Oncol Biol Phys* vol.13, no.2, pp283-287, February, 1987.

[11] B. Jones and K.I Hopkins, "Chemotherapy and other adjuvant therapies," in *Radiobiological modelling in radiation oncology*, R.G. Dale and B. Jones, Ed. London: British Institute of Radiology, 2007,pp234-345.

[12] V. Grégoire and M. Baumann. "Combined radiotherapy and chemotherapy," in *Basic Clinical Radiobiology*, M. Joiner and A. van der Kogel, Ed. London: Hodder Arnold, 2009, pp246-258.

[13] A. Cárabe-Fernández, R.G. Dale, J.W Hopewell, B. Jones and H. Paganetti, "Fractionation effects in particle radiotherapy: implications for hypo-fractionation regimes," *Phys Med Biol* vol. 55, no. 19, 5685-5700,October, 2010.

[14] G.A. Plataniotis and R.G. Dale "Use of the concept of equivalent biologically effective dose (BED) to quantify the contribution of hyperthermia to local tumor control in radiohyperthermioa cervical cancer trials, and comparison with radiochemotherapy results," *Int J Radiat Oncol Biol Phys* vol. 73 no. 5, pp1538-1544, 2009.

[15] [15]G.S. Stamatakos, E.A. Kolokotroni, D.D. Dionysiou, E.C. Georgiadi and C. Desmedt, "An advanced state-discrete event multiscale simulation model of the response of a solid tumor to chemotherapy: Mimicking a clinical study," *J Theor Biol* vol.266, no. 1, pp 124-139, May, 2010.

Modular Markup for Simulating Vascular Tumour Growth*

David Johnson, Anthony J. Connor, and Steve McKeever

Abstract— **In this paper we present preliminary results of the first application of TumorML being developed outside of the context of the Transatlantic TUmor Model Repositories project (TUMOR). Based on a domain-specific software framework for developing models to simulate vascular tumour growth, we have developed a corresponding domain-specific language (DSL) for use with the framework. The DSL script is directly embedded into TumorML model descriptions serving as an example of how within a single model description document, we can fully describe cancer models as functional components. We introduce the framework that our DSL orchestrates; show fragments of DSL script we have developed to describe tumour-induced angiogenesis; and how these functional model descriptions can be integrated and executed with TumorML markup.**

I. Introduction

In computational biology, there is a diverse range of programming and descriptive languages that span across different biological domains and scales. This creates challenges for model reuse and composition, since each model implementation, even if available, may use a completely different technological framework. Combining models may therefore require porting models to a new framework, or re-implementing them, both costly and error prone activities. Before the year 2000, there were no unified efforts towards standardized languages for describing models. Markup languages for computational biology emerged soon after the turn of the millennium with the SBML (Systems Biology Markup Language) [1] and CellML [2] research programmes.

Generally speaking, all application-specific markup languages are based on the eXtensible Markup Language (XML) [3], as are SBML and CellML. XML emerged as a popular choice for computer-based language definition in the late 1990's as it defines a standard syntax on to which other vocabularies can be built. This allows language specific parsers to reuse the standard XML parsing routines for processing XML documents. XML-based languages are therefore well suited as software-neutral information exchange formats. XML can be thought of as the base alphabet and grammar of a language. What raw XML lacks is semantics or definitions to provide context and application. This is provided in part by specific vocabularies built on XML, which define element and attribute names and the structural relationships between them. Multiple XML vocabularies can be combined within a single document, enabling the development of various languages targeted at specific narrow domains of discourse that can be incorporated into a compound language.

We have previously discussed in [4] the drawbacks of the current state-of-the-art in SBML, CellML and insilicoML [5], where we do not believe any is expressive enough to capture the multi-scale nature of cancer models, and diverse techniques typically involved in the development of a cancer model such as ODEs, and PDEs. The Transatlantic TUmor MOdel Repositories project (TUMOR) aimed to enable interoperation between cancer model repositories, and to fulfil this aim we developed TumorML, a markup language for *in silico* oncology, first mentioned in [6] and early groundwork described in [7]. The original intent was to enable interoperation between parts of the infrastructure developed under TUMOR, specifically to enable model exchange between repositories in the United States' Center for the development of a Virtual Tumor (CViT) [8] and Europe. However, in the longer term we view TumorML as a markup standard with wider applications than that of the original remit of TUMOR.

Initially, TumorML was envisaged as a markup language that wraps existing software implementations of models in metadata to curate models in digital repositories, and to enable execution of models on computational frameworks. As such, the functional description of a model is treated as a proverbial 'black box' – how a model is implemented is not of concern in the model description; only metadata pertaining to parametric interfaces (i.e. I/O of data) and the minimal hardware and software execution requirements are ever described. The rationale for designing TumorML in this fashion was to enable support for existing implementations of models where the models directly concerning TUMOR (developed out of the ACGT and ContraCancrum projects) had already been implemented and optimised in the C and C++ general-purpose programming languages (GPLs). The work described in this paper addresses two drawbacks of this approach of wrapping up existing implementations with markup. Firstly, model implementations are difficult to validate where the underlying mathematics and algorithms used are not exposed. Both automated and human validation therefore cannot be carried out if model codes are not available for inspection. Secondly, implementations developed using GPLs are more susceptible to security vulnerabilities, as no restrictions to I/O or memory access are imposed. In theory, any executable tailored for any purpose can be wrapped in TumorML markup.

Constraining the descriptive domain is one way to address these drawbacks, and domain-specific languages (DSLs) can be used for model description. With a more restrictive description language, we can more easily validate the methods used. By only making available certain I/O to the model developer (i.e. being able to read and write a certain set of data types) we may be able to limit erroneous and

*Contributions by DJ and SM in this paper are financially supported in part by the EC under the Transatlantic TUmor MOdel Repositories project (#FP7-ICT-2009.5.4-247754). Contributions by AJC are supported in part by the UK Engineering and Physical Sciences Research Council and Hoffman la-Roche Ltd.

David Johnson is with the Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford, OX1 3QD, United Kingdom (phone: +44 (0) 1865 273838; fax: +44 (0) 1865 273839; e-mail: david.johnson@cs.ox.ac.uk).

Anthony J. Connor is with the Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford, OX1 3QD, United Kingdom (e-mail: anthony.connor@keble.ox.ac.uk).

Steve McKeever is with the Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford, OX1 3QD, United Kingdom (e-mail: steve.mckeever@cs.ox.ac.uk).

malicious exploitation of model code. The rest of this paper describes a domain-specific framework that is based on the reverse engineering of a class of models developed by Alarcón et al, looking specifically at how tumours stimulate vascular growth to enable proliferation of cancerous cells. A DSL is presented to orchestrate the framework, and we describe how to combine our DSL together with TumorML.

II. Simulating Vascular Tumour Growth

Vascularisation of a tumour marks the transition of that tumour from being essentially harmless to increasingly invasive and eventually fatal [9]. Tumours during the avascular growth stage are typically manageable and relatively harmless. They are limited in size due to the lack of nutrients available to fuel their proliferation. However, once a tumour grows too large to be supported by the oxygen supplied from existing vasculature, the tumour cells secrete various angiogenic growth factors, including tumour angiogenesis factors (TAFs). TAFs diffuse through the healthy tissue surrounding the tumour and upon reaching a blood vessel will stimulate vascular growth towards the tumour. Over time the vascular network evolves to a state in which the level of nutrients and oxygen in the tissue surrounding the tumour have increased sufficiently to allow for further growth of the tumour. Multiscale models are commonplace in biological modelling, as they also are within the domain of cancer modelling. Modelling tumour growth involves the integration of a number of biological models concerning different processes as part of a bigger picture. Each model may also integrate different mathematical techniques to adequately model these processes [10].

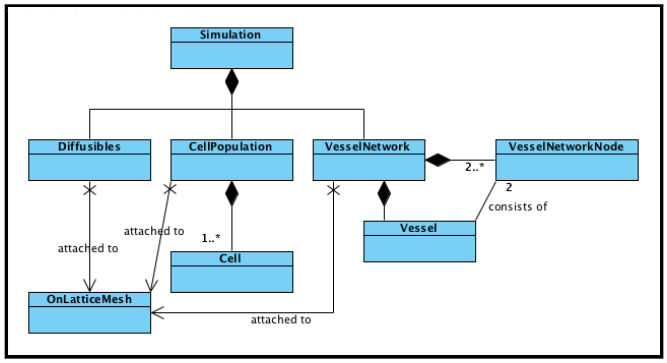


Figure 1. UML class diagram illustrating the OO framework's key complex data structures contained within the `Simulation` controller object.

The University of Oxford's Department of Computer Science, in conjunction with the Centre for Collaborative Applied Mathematics and the Centre for Mathematical Biology, is actively developing an object-oriented (OO) modelling framework for implementing hybrid and multiscale models of vascular tumour growth. The focus of the framework development is to apply software engineering techniques that allow it to be highly reusable and extensible – characteristics that have, to the best of our knowledge, eluded the majority of biological modelling software frameworks to date.

Work is underway to reverse-engineer a family of models published by Tomás Alarcón, Helen Byrne, Philip Maini and collaborators, in particular the family of models discussed and extended in [14], to extract and abstract the common methodologies and data structures involved in the development of simulations based on vascular tumour growth models. The OO framework has been developed based on these abstractions, and a functioning implementation of the framework is being developed in C++. The framework is presented fully by Anthony J. Connor et al in a forthcoming paper [16].

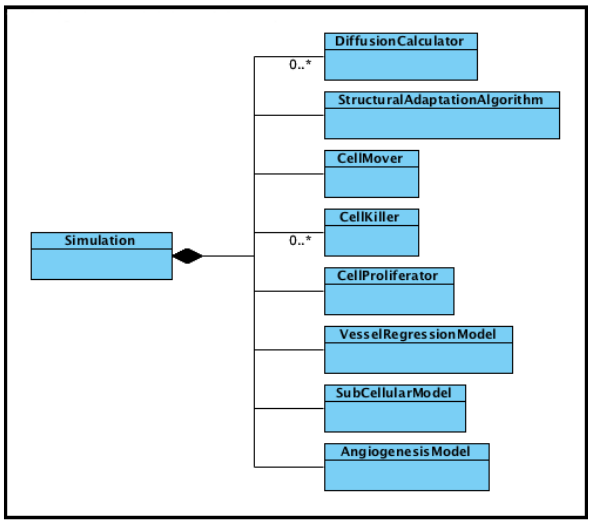


Figure 2. UML class diagram illustrating the containment hierarchy of functional aspects within the OO framework.

The OO framework consists of a set of commonly used complex data structures modelling relevant biological entities (Figure 1) and functional aspects that allow modular coupling of different algorithms (Figure 2) by employing the *visitor* and *strategy* design patterns [11], [12]. These design patterns decouple the biological entities from the algorithms that act on them. This allows algorithm code to be easily replaced without modifying the framework itself.

Figure 1 shows the static `Simulation` class containing a `Diffusibles` object that holds a list of diffusible chemicals within the simulated environment. A `CellPopulation` object contains a collection of `Cell` objects to represent the population of normal and cancerous cells. Finally, a `VesselNetwork` object models the vascular network as a graph of `Vessel` objects. Each of these is attached to a lattice-based mesh structure, `OnLatticeMesh`. Figure 2 illustrates each of the different abstract algorithm or model container classes. These are subclassed with concrete implementations by model developers. The implementations are substituted for these abstract classes at runtime. The `Simulation` class acts as a controller with an encapsulated algorithm that behaves according to each of the component classes' implemented code (i.e. based on code substituted for `DiffusionCalculator`, `CellMover`, `CellKiller` etc.).

Developing a framework with common abstractions of biological entities, and decoupling entities from algorithms

that model the different biological processes brings two key advantages. Firstly, it provides model developers a modular environment in which submodels and individual algorithms can be tested in isolation before being integrated into larger models of vascular tumour growth. Secondly, taking an OO approach to abstracting common structures and processes of vascular tumour growth models should make models more understandable for biologists and mathematical modellers who are not well versed in computer programming.

III. Modular Markup using TumorML

We are currently working with modellers to develop the next level of more detailed abstraction in the structural, mathematical, and algorithmic descriptions of the inner workings of models. While significant effort might be required to port existing models to TumorML, providing multiple levels of abstractive notation in our markup allows us to wrap existing models in early versions of TumorML as well as publish new models with an evolving markup specification. Version 1.0 of the TumorML specification is publicly available from the TUMOR website (www.tumor-project.eu) and the latest XML schema is available open-source from the SourceForge website (www.sf.net/p/tumorml).

To support our OO framework for modelling vascular tumour growth, we are developing a DSL that can be cross-compiled into C++ code compatible with the implementation of our framework. Scientists will implement their models within the constraints of the framework using the DSL script. This will bring several benefits. Adopting a standard view of the most common abstractions (e.g. cells, populations of cells, intracellular chemicals), the most typical behaviours of this class of model (e.g. moving cells, killing cells, subcellular biochemical processes), and I/O using standard formats, means that different models developed within the framework can more easily be validated. Limiting the functionality of the DSL also reduces the risk of malicious code being introduced into model codes.

```
Alarcon05OxygenDependentModel extends SubCellularModel {
   ...
   // When p53 concentration exceeds threshold, kill cell.
   if (Cell.getChemicalConcentration(Chemical.p53) >
Parameters.normalCellDeathThresholdConcentration)
      Cell.kill();
   // When cycle time exceeds divide time, flag cell to divide.
   if (Cell.cellCycleTime >= Parameters.cellDivideTime) {
      // Flag cell to divide.
      Cell.divide();
      // Reset the cell cycle time of this cell.
      Cell.setCellCycleTime(Cell.cellCycleTime -
Parameters.cellDivideTime);
      }
   }
}
```

Figure 3. A DSL script fragment illustrating how our OO framework extends the `SubCellularModel` abstract class allowing model developers to provide their own concrete implementation.

Figure 3 illustrates how to implement a new subcellular model within the OO framework. This model at runtime is substituted where the `SubCellularModel` class is defined within the containment hierarchy. In this example, a model by Alarcón et al described in [13] is used as a determinant for the subcellular conditions for cell death and proliferation. The script encapsulated here is executed at each time step for each `Cell` object in the cell population, as determined by the `Simulation` controller.

```
...
Simulation.setSubCellularModel(Alarcon05OxygenDependentModel);
Simulation.setDuration(Parameters.simulationRunTime);
Simulation.setTimestep(Parameters.timestep);

// Create the lattice and add some diffusables chemicals to the
environment.
l = Lattice(Parameters.siteSize, Parameters.domainSize_x,
Parameters.domainSize_y);
d = Diffusables(lattice);
d.addSpecies(Chemical.VEGF);
d.addSpecies(Chemical.Oxygen);
...
Simulation.run(); // go!
```

Figure 4. DSL code fragment shows how a subcellular model, such as that illustrated in figure 3, can be attached to a simulation.

To show how we can augment DSL script with XML markup, we have extended the TumorML schema to allow the scripts to be embedded into TumorML's implementation element within `<script>` tags, replacing the XML elements that are normally used for publishing conventional source code and pre-compiled binary files.

```
<tumorml xmlns="http://www.cs.ox.ac.uk/tumorml/1.0s" id="Al05>
   <header>
      <title>Vascular Tumour Growth Model Example</title>
      ...
   </header>
   <model>
      <parameters>
         <in name="simulationRunTime" optional="false">
            <value type="int" unit="second" />
         </in>
         ...
      </parameters>
      <implementation>
          <script>
             ...
             Simulation.setSubCellularModel(Alarcon05OxygenD...
             Simulation.setDuration(Parameters.simulationRun...
             ...
             Simulation.run();
          </script>
      </implementation>
   </model>
</tumorml>
```

Figure 5. An example of DSL script embedded within a TumorML document, illustrating parameter mapping from the XML declaration to the script reference for the input parameter `simulationRunTime`. Note some elements have been omitted for clarity.

For accessing externalized parameters, the framework utilises TumorML's parameter definitions. The framework reads and parses the XML parameter definitions that describe the inputs and outputs for computational models and builds an in-memory parameter map, where referencing parameters using string-based labels retrieves values that can be inserted into a model's script. This provides us with a standard method of executing model code when using a variable range of parameter values. The TumorML parameter definition also includes metadata describing the command-line data type, quantity/unit type, and optionally an identifier referencing an externally accessed biological ontology such as SBO, the Systems Biology Ontology (www.ebi.ac.uk/sbo/), or BioPAX, Biological Pathway Exchange (www.biopax.org). This gives us a mechanism by which pre-runtime checks can be carried out, as well as enabling intelligent checks to be performed when combining code modules and providing extra human readable information to model authoring,

workflow composition, and clinical diagnosis or research tools.

Illustrated by Figure 5, the static `Simulation` object is set up with any predefined settings through hard-coded default values or linked externally via a parameters definition such as that in TumorML whereby they are accessed in the DSL script with the `Parameters` keyword. In this example, the XML input parameter given the name `simulationRunTime` maps directly to the DSL script referenced by `Parameters.simulationRunTime`. Variables can also be declared for brevity, as shown with the `Lattice` and `Diffusibles` objects referenced by `l` and `d` respectively. The `Chemical` keyword is used to refer to a statically defined enumeration of chemicals, including TAFs. In this example we make reference to the vascular endothelial growth factor (`Chemical.VEGF`) protein and oxygen (`Chemical.Oxygen`). The mathematics and algorithms used in the framework are described with syntax similar to that of ECMAScript [15], which many of the scripting conventions in our DSL are based on. As our DSL is at an experimental stage, a full specification will be published at a later time.

IV. Results

We have demonstrated how domain-specific markup can link transparently with model implementations by describing vascular tumour growth with our own defined DSL. Our use-case driven design process has yielded a DSL that includes conventional programming constructs, modern OO methodologies, including design patterns, and access to numerical libraries.

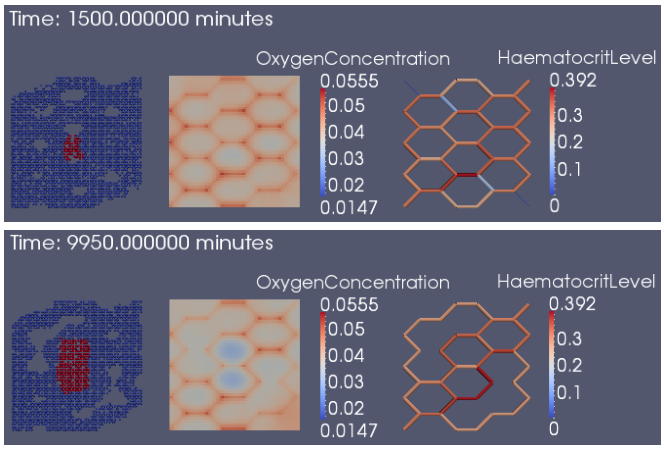


Figure 6. A pair of snapshots showing a two-dimensional simulation of vascular tumour growth generated from our OO framework.

At present, our OO framework can produce simulations of vascular tumour growth where we reproduce the family of models that the work in [14] builds on. Figure 6 shows two snapshots taken from a simulation implemented within our framework coupling a number of models including an oxygen dependant model described in [13] by Alarcón et al. On the left hand side cell populations are visualized where cancerous cells are depicted in red and healthy cells in blue. The centre panel shows the oxygen distribution. The right hand side illustrates the evolving vessel network with vessel radii and hematocrit (packed cell volume or erythrocyte volume fraction) levels. The models that produce these simulations can be written in our DSL and wrapped up in TumorML markup. The models are then cross-compiled to an efficient C++ implementation using the OO framework.

V. Conclusions

By introducing a DSL into TumorML markup, we have shown how domain-specific markup can link transparently with model codes to describe models of vascular tumour growth. At present, our OO framework can reproduce simulations of vascular tumour growth such as that pictured in Figure 6. In this paper we have described the OO framework being developed for modelling vascular tumour growth, and present initial work on a DSL for orchestrating the framework. Integrating our DSL with TumorML allows us to package models as modules, which is essential for enabling multiscale model composition and reuse, and for publishing to repositories such as the TUMOR model repository or the CViT Digital Model Repository. Although using a domain-specific framework limits the scope of what can be modelled, we submit that wholistic description and transparency is key to models being adopted, reused, and further developed by the cancer modelling community.

References

[1] M. Hucka et al, “Evolving a lingua franca and associated software infrastructure for computational systems biology: the Systems Biology Markup Language (SBML) project,” *Systems Biology*, vol. 1, no. 1, pp. 41-53, Jun. 2004.

[2] C.M. Lloyd et al, “CellML: its future, present and past,” *Progress in Biophysics and Molecular Biology*, vol. 85, pp. 433-450, Jun. 2004.

[3] T. Bray et al (eds.), *Extensible Markup Language (XML) 1.0 (Fifth Edition) W3C Recommendation* [Online]. Available: www.w3.org/TR/2008/REC-xml-20081126/

[4] D. Johnson et al, “Markup languages for in silico oncology,” in *Proc. 4th Int. Adv. Res. Workshop on In Silico Oncology and Cancer Investigation (4th IARWISOCI) The ContraCancrum Workshop*, Athens, Greece, 2010, pp. 108-110.

[5] Y. Asai et al., “Specifications of insilicoML 1.0 : a multilevel biophysical model description language.” *J. Physiological Sci.*, vol. 58, no. 7, pp. 447–458, Dec. 2008.

[6] D. Johnson and S. McKeever, “The case for using markup for biomechanical modelling,” in *Proc. Int. Soc. Biomechanics XXIII World Congress*, Brussels, Belgium, 2011.

[7] D. Johnson et al, “TumorML: concept and requirements of an in silico cancer modelling markup language,” in *Proc. 33rd Annu. Int. Conf. IEEE EMBS*, Boston, MA, 2011, pp. 441-444.

[8] T.S. Deisboeck et al., “Advancing cancer systems biology: Introducing the Center for the Development of a Virtual Tumor, CViT.” *Cancer Informatics*, vol. 5, pp. 1–8, Mar. 2007.

[9] J. Folkman, “Tumor angiogenesis: therapeutic implications.” *New England Journal of Medicine (NEJM)*, vol. 285, no. 21, pp. 1182–1186, Nov. 1971.

[10] T.S. Deisboeck et al, “Multiscale cancer modelling.” *Annu. Review Biomedical Eng.*, vol. 13, pp. 127–155, 2011.

[11] B. Mayer and K. Arnout, “Componentization: the visitor example.” *IEEE Computer*, vol. 39, no. 7, pp. 23-30, Jul. 2006.

[12] E. Gamma et al., *Design Patterns, Elements of Reusable Object-Oriented Software*. Addison-Wesley, 1995.

[13] T. Alarcón et al, “A multiple scale model for tumor growth.” *Multiscale Model Simulation*, vol. 3, no. 2, pp. 440–475, 2005.

[14] M.R. Owen et al, “Angiogenesis and vascular remodelling in normal and cancerous tissues.” *J. Math. Biology*, vol. 58, no. 4-5, pp. 689-721, 2009.

[15] *ECMAScript Language Specification, Edition 5.1*, ECMA-262, June 2011.

[16] A.J. Connor et al, “Object oriented paradigms for vascular tumour growth: a case study,” unpublished, submitted to IARIA SIMUL 2012.

The Health Data Ontology Trunk (HDOT). Towards an Ontological Representation of Cancer-related Knowledge*

Emilio M. Sanfilippo, Ulf Schwarz, and Luc Schneider

***Abstract* – Ontologies play a central role in new IT strategies for personalized medicine that enhance the communication between patients and health care providers. In this paper we present the Heath Data Ontology Trunk (HDOT) and how it is used for this purpose within the p-medicine project. We describe crucial methods in its development, in particular its modularity. The main benefit of a domain middle-layer ontology approach is that the latter can be further extended and specialized in different but interrelated modules according to more specific biomedical and clinical needs and requirements within the same semantic framework. As an example of a modular extension of HDOT relevant to cancer research and treatment we describe a first prototype of the HDOT pathological formation module. We demonstrate the usefulness of this module in two scenarios: the semantic integration of heterogeneous biobank metadata and the design and administration of Case Report Forms (CRFs).**

I. P-MEDICINE'S SEMANTIC STRATEGY

Medicine is undergoing a revolution that is transforming the nature of healthcare from reactive to preventive and to a personalized predictive treatment. This transformation process has motivated the concept of the Virtual Physiological Human[1] (VPH) that seeks to develop a scientific methodological and technological framework to enable the construction of models of the human body as a single complex dynamical system. Nevertheless, there is still the need for the development of a sound multidimensional modeling of the natural phenomenon of cancer supporting both clinical investigations and the emergent scientific, technological and medical discipline of *in silico oncology* [1]. 'p-medicine: From data sharing and integration via VPH models to personalized medicine'[2] is a four-year Integrated Project co-founded under the European Community's 7th Framework Programme which proposes to create an infrastructure that will facilitate the development from current medical practice to personalized medicine, with particular attention on cancer-related information.

In p-medicine special attention is paid to patients who are not just treated as passive consumers of health care services, but are continuously and actively involved in the comprehension of their health status and management of their clinical trials. Data integration, standardization and semantic interoperability in p-medicine are ontology-driven. Data and information coming from different sources (mainly clinical care providers, both external and internal to p-medicine, clinical trial centers, laboratories, research institutions and bio-banks) are pushed into the p-medicine warehouse and at the same time annotation are generated by an annotation service using the Health Data Ontology Trunk (HDOT). The original data together with a mapping file are then stored in the p-medicine data warehouse; the integrated data is converted to RDF format and stored in a triple store so that it can be easily exploited later for different purposes according to users' needs.

II. THE HEALTH DATA ONTOLOGY TRUNK (HDOT)

The Health Data Ontology Trunk (HDOT) is being developed by the Institute for Formal Ontology and Medical Information Science (IFOMIS) of the University of Saarland and it is conceived as a modular middle-layer ontology. HDOT is being designed, maintained and extended using the ontology editor Protégé[3], and is released in OWL-DL[4] under the following web address: http://code.google.com/p/hdot/[5].

HDOT integrates under the same semantic umbrella (figure 1) the first version of the Basic Formal Ontology[6] (BFO 1.1), the Relational Ontology[7] (RO), the Information Artifact Ontology[8] (IAO), the Middle

* Research supported by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No 270089 p-medicine, as well by the National Research Fund, Luxembourg, with the support of the Marie Curie Actions of the European Commission (FP7-COFUND).

Emilio M.Sanfilippo is with the Institute for Formal Ontology and Medical Information Science (IFOMIS) of the University of Saarland, PO- Box. 151150, D-66041 Saarbruecken, Germany (phone: 004968130264775, fax: 004968130264772, e-mail: emilio.sanfilippo@ifomis.uni-saarland.de).

Ulf Schwarz is with the Institute for Formal Ontology and Medical Information Science (address above) (phone: 004968130264770, e-mail: ulf.schwarz@ifomis.uni-saarland.de).

Luc Schneider is with the Institute for Formal Ontology and Medical Information Science (address above) (phone: 004968130264773, e-mail: luc.schneider@ifomis.uni-saarland.de).

[1] http://www.vph-noe.eu/

[2] http://www.p-medicine.eu/

[3] http://protege.stanford.edu/

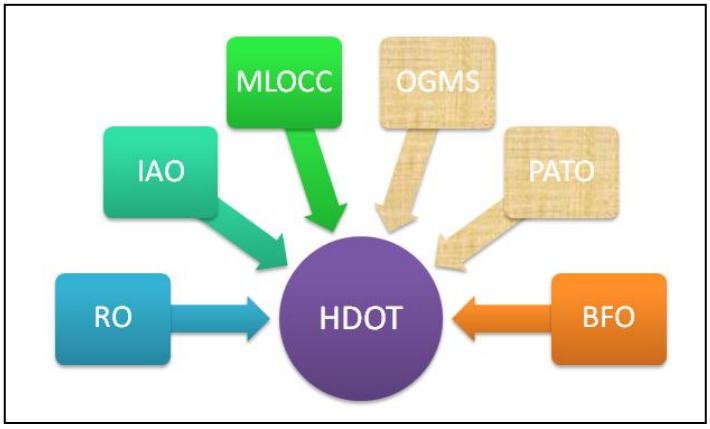
[4] Description logic expressivity: SROIQ (D)

[5] To explore HDOT and its modules using Protégé, you have to open the google code page, go to the page "source" and then to "browse". Click on "trunk" on the right column; open the file you want to explore (hdot.owl for example), open the link "view raw file" on the right side of the page with Protégé.

[6] http://obofoundry.org/cgi-bin/detail.cgi?id=bfo. BFO is currently undergoing an update process to BFO 2.0. However, a stable release of BFO 2.0 has not been released yet

[7] http://obofoundry.org/cgi-bin/detail.cgi?id=relationship

[8] http://obofoundry.org/cgi-bin/detail.cgi?id=information_artifact



Layer Ontology for Clinical Care[9] (MLOCC) [2], parts of the Phenotypic Quality Ontology[10] (PATO) and parts of the Ontology for General Medical Science[11] (OGMS).

Most of them are part of the OBOFoundry initiative [3] and are widely used in the biomedical domain for data annotation and integration.

HDOT is designed in a modular fashion [4] as a middle-layer ontology in the sense that it specifies upper-level domain independent classes down to the biomedical domain while maintaining at the same time a very general semantic and axiomatic structure that can be further developed and specialized in different modules for different purposes and applications. Therefore a middle-layer ontology can be thought as an ontology whose classes are domain-driven but application-independent, in the sense that they represent general structural characteristics and properties of the domain without specializing them in very specific semantic and axiomatic features. HDOT is intentionally developed, maintained and further extended as a modular ontology, so that we have the possibility to add more information and introduce new subclasses according to users' needs without ever having to alter its overall structure. One core requirement put on HDOT is that it is broad enough to contain all general classes under which all necessary and more specific semantic content can be subsumed in such a way that a meaningful axiomatic relational structure is conferred upon those specific area of contents. Indeed, we can enrich the semantic content of different resources by subsumption under classes of HDOT, e.g. integrating specific parts of biomedical terminologies like ICD-10, the NCI thesaurus or SNOMED-CT, in order to enhance their formal and semantic constraints in a computable way.

HDOT's development as a middle-layer ontology is governed by three main related structural considerations in order to achieve the highest level of semantic interoperability between heterogeneous data sources, maintain a high level of ontological soundness and ensure a high degree of expandability:

1. The HDOT level of generality is designed in such a way that HDOT classes and relations cover all areas of the health-care domain, i.e. there is a meaningful ontologically well-defined HDOT super-class under which all necessary parts and pieces of semantic data descriptions (annotations, metadata) can be directly subsumed or otherwise represented[12]. The semantics of HDOT central body is supposed to change in the further development of the project only in case problems related to HDOT's application to the project itself or clinicians' needs emerge;

2. The core ontological structure integrates different modular ontologies at different levels of granularity. Each class is provided with a deep axiomatization, which guarantees to the users' work-flow high degrees of semantic representation and syntactic reasoning, together with the ability to construct defined classes and composite terms.

3. HDOT's modules for specific applications can be obtained by stating further specifications of HDOT classes, i.e. by inserting subclasses under existing HDOT slots (super-classes). We thus accord with Parent and Spaccapietra [5] who consider an ontological module to be a smaller portion of an ontology that is intended to be a subset of a broader knowledge organization.

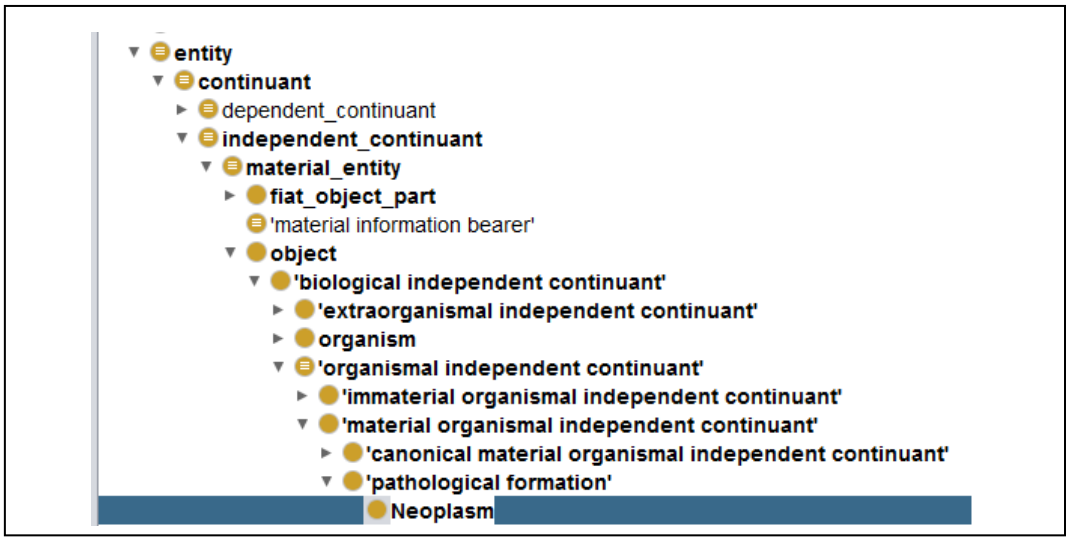
The most important owl-class axioms are governed by the ordering relation of subsumption between classes while most of the other axioms originate from the following four considerations: (1) To provide machine readable and computable class constructions; (2) to provide ontologically sound relations between classes and the corresponding labels to enable the desired reasoning and inference capabilities; (3) to provide the basis for the composition of biomedical complex concepts (e.g. 'blood pressure increase', 'resection of tumour in kidney') by axiomatizing the

[9] http://www.ifomis.org/chronious/mlocc

[10] http://obofoundry.org/cgi-bin/detail.cgi?id=quality

[11] http://obofoundry.org/cgi-bin/detail.cgi?id=OGMS

[12] However, it is worth noting that HDOT is currently under development, therefore it is very probable that the domain coverage is not yet complete.



necessary relations between HDOT's constituents; (4) to include relations which bridge different levels of granularity (e.g. "part_of" or "contains").

Since p-medicine is presently focused on the computational representation of data about neoplasms, we extended HDOT with a first module particularly suitable for the representation of cancer-related information.

III. PATHOLOGICAL FORMATION MODULE UNDER HDOT

The HDOT pathological formation module specifies in more details information about cancer integrating the ACGT Master Ontology (ACGT MO) [6]. We decided to integrate this ontology for four main reasons: 1) its domain coverage is very close to HDOT; 2) its axiomatic structure is very similar to HDOT; 3) it had undergone an evaluation process that is very well documented [7]; 4) p-medicine is accounted as the follow-up of the ACGT project[13].

Following the ontological approach of OGMS and specifying its structure according to our needs, we reused the class OGMS_0000077 'pathological formation' and specialized it with the ACGT MO subclass 'neoplasm'. We adhered to the biomedical specification of BFO's class 'material entity' throughout this extension process (figure 2), which was possible because both OGMS and ACGT MO are built on BFO.

From an axiomatic point of view, we thus consider a pathological formation in HDOT as an enduring material entity which is clinically abnormal and is the outcome of a pathological process which occurs in a particular region the human body and realizes a dispositional disease such as cancer [8].

Therefore we need a three-fold ontological representation:

- The disease (e.g. cancer) as a BFO disposition;
- The pathological process as a BFO occurent which in this case realizes a disease;
- The pathological formation as a BFO material entity as an outcome of the pathological process (e.g. neoplasm).

In owl-axioms these levels can be described in the following form:

PathologicalFormation *isSubclassOf* **some** MaterialEntity **and** *isOutcomeOf* **some** (PathologicalProcess **and** (*realizes* **some** Disposition)),

where a pathological formation, as well as a pathological process are regarded as *clinically abnormal* in the sense that they are not part of the life plan of an organism of the relevant type (unlike aging, pregnancy or menopause) [8].

IV. TWO SCENARIOS

A. *Mapping metadata of biobanks to hdot*

Biobanks represent key resources for clinico-genomic and personalized medical research. It is crucial that scientists can securely access and share high quality information on biomaterial and related metadata, resulting in the need to integrate biobanks into larger biomedical ICT infrastructures. This integration presupposes semantic interoperability between the metadata used in the different biobank information systems. This cannot be achieved by lexical mappings only because even if one and the same term or expression occurs in different biobank metadata sets, there is no guarantee that this term is used coherently in the same sense.

Therefore, semantic interoperability between heterogeneous biobank metadata can be assured by mapping the latter onto a biomedical middle-layer ontology such as HDOT. Its semantic and axiomatic framework would help scientists to search for and request information on biomaterial more accurately for their specific research projects, without running the risk of semantic and axiomatic incompatibilities between their different knowledge bases. In addition,

[13] http://eu-acgt.org/

HDOT could allow biobank owners to share their biomaterial and corresponding data stored in their databases within the same ontological structure, bringing the benefits of semantic integration among the whole scientific community.

B. *Pre-coordinated annotation of data from case report forms (CRFs)*

A similar data integration problem occurs in the design and administration of case report forms (CRFs). They come from heterogeneous sources and are used to collect a plethora of information. Data gathered in CRFs is often annotated after the collection process and rarely a common semantic framework is used. The exploitation of HDOT already in the design phase of a CRF can ensure adequate data annotations for the collected data in a pre-coordinated way. We showed elsewhere [9] that CRF questions can be generated directly from ontological axioms utilizing relational paths in an ontology.

V. OTHER ONTOLOGICAL APPROACHES

In the current state of the art of ontological engineering for the biomedical domain, there are several resources representing data about cancer and tumour. The BioPortal[14], a service maintained by the US National Center for Biomedical Ontologies to access freely biomedical ontologies, contains 326 ontologies and 5,232,009 terms[15]. Looking for 'neoplasm', it retrieves 145 results in 24 resources[16]. Among these, only 10 are proper ontologies in strict sense [3], while the others are terminologies, coding systems and thesauri. Unfortunately, most of the latter suffer from deep ontological flaws such as blurring important distinctions between different real world entities, clinical findings and diagnosis, as well as between biological processes and dispositional entities. Thus, we argue that using BioPortal services at face value and without further processing and elaboration can seriously infringe the ontological soundness of semantic data integration approaches[17].

VI. CONCLUSIONS

Personalized medicine is currently facing the problem of making patients directly involved in their health care. New IT strategies are currently explored to enhance the communication between patients and health care providers. Ontologies play a central role in this framework, because they help in many ways, such as making people's assumptions explicit and representing the complexity of the biomedical domain in a computable language, thus making data sharing more efficient. In this paper we present the Health Data Ontology Trunk (HDOT). We describe crucial methods in its development in particular its modularity and the benefits of a middle-layer ontological approach. Indeed, a domain middle-layer ontology like HDOT can be further extended and specialized in different but interrelated modules according to more specific biomedical and clinical needs and requirements while always maintaining one and the same axiomatic framework. One major advantage of this approach is that we can specify necessary top-class axioms in the middle-layer that are automatically inherited by appended subclasses. In this way we provide a minimal set of semantic constraints on the use of biomedical and clinical terms and expressions. As an example of a modular extension of HDOT we present a first prototype of the HDOT pathological formation module whose semantics and axiomatic structure are particularly focused on the ontological representation of cancer-related information. We demonstrate the usefulness of our approach in two scenarios.

[14] http://bioportal.bioontology.org/

[15] BioPortal statistics, from the BioPortal website, last access on 09.10.2012

[16] http://bioportal.bioontology.org/search?query=neoplasm&commit=Search, last access on 09.10.2012

[17] Projects like VPH Share are currently facing these issues

REFERENCES

[1] N. Graf.et al., "In silico oncology for clinical decision making in the context of nephroblastoma", Klin Padiatr. Vol. 221(3), pp.141-92009, May-Jun 2009;

[2] M. Brochhausen, L. Schneider, "The CHRONIOUS Ontology Suite: Methodology and Design Principles". In Smith Barry et al. (eds.): Proceedings of the International Conference on Biomedical Ontologies (ICBO 2011), SUNY University Buffalo, Buffalo/NY, 2011, 167-173., available from http://ceur-ws.org/Vol-833/paper22.pdf

[3] B. Smith et al., "The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration", Nature Biotechnology vol. 25, pp. 1251 – 1255, 2007

[4] H. Stuckenschmidt, et al. (eds.), Modular Ontologies. Concepts, Theories and Techniques for Knowledge Modularization, Berlin-Heidelberg: Springer-Verlag, 2009

[5] C. Parent, S. Spaccapietra, An Onverview of Modularity, in Stuckenschmidt [2009], pp.5-23

[6] C. Cocos, "ACGT D7.7 Design Principles of the ACGT Master Ontology: Examples and Discussion", available from: http://www.ifomis.org/projects/wiki/images/7/7e/ACGT_D7.7_IFOMIS_Final.pdf

[7] G.Gintarė et al., "Evaluating Ontologies with NLP-Based Terminologies – A Case Study on ACGT and Its Master Ontology", in Formal Ontology and Information System (FOIS). Proceedings of the Sixth International Conference, Amsterdam, IOS Press Amsterdam, pp.331-342

[8] R.H. Scheuermann et al., "Toward an Ontological Treatment of Disease and Diagnosis", Proceedings of the 2009 AMIA Summit on Translational Bioinformatics, San Francisco, CA, 2009, pp- 116-120

[9] M. Brochhausen, et al., "The ACGT Master Ontology and its Applications – Towards an Ontology-Driven Cancer Research and Management System", available from http://www.ifomis.org/projects/wiki/images/8/85/BrochhausenEtAl.pdf

Scientific Workflows to Support *In Silico* Modeling in Cancer Research*

Stelios Sfakianakis, Vangelis Sakkalis, and Kostas Marias, *Member, IEEE*

***Abstract*— In order to deal with the inherent complexity in cancer new computational models are being developed to aid both biological discovery and clinical medicine. The development of these in silico models is facilitated by rapidly advancing experimental and analytical tools that generate information-rich, high-throughput biological data. In this paper we propose the use of the scientific workflows for the integration of multilevel in silico models and we describe a technical infrastructure built on state of the art web technologies for the implementation of these workflows.**

I. Introduction

Cancer is an intrinsically complex and heterogeneous disease, making it particularly amenable to systems biology approaches [1]. Computational biomodelling has emerged as a new field in the crossroads of molecular biology, biochemistry, computer science, and mathematics to provide a better understanding and simulation of key biological processes in living organisms [2]. So far, significant but highly fragmented efforts have been made on both sides of the Atlantic to develop and use models of pathophysiology in order to better understand human function and promote individualized, patient-specific optimization of cancer treatment. The TUMOR project (http://tumor-project.eu), an EU FP7 funded project, is developing a European clinically oriented semantic-layered cancer digital model repository from existing EU Virtual Physiological Human (VPH) related projects designed to be interoperable with the US grid enabled semantic-layered digital model repository platform at CViT which is NIH/NCI-caGRID compatible. Models and data will drive advances in cancer modeling with the ultimate goal to build an integrated, interoperable transatlantic research environment offering the best available models and tools for clinically oriented cancer modeling and serving as an international validation/ clinical translation platform for predictive, in silico oncology.

The main users of such a system are the VPH modelers and biomedical researchers that are focusing on the comprehension of the biological processes and interactions in nature and their simulations through computational models. The primary concerns of these users are:

- To share and reuse biological and physiological models.
- To share and reuse biomedical data.
- To plug (combine) models together to create larger, more comprehensive models without excessive demands for user input.
- To execute (run) the simulation models, trace their execution, and keep an archive of the results and the history of these executions.

In order to simulate cancer growth and simulate the efficacy of specific treatments in a more holistic way different models with diverse biocomplexity levels and directions (bottom-up, top-down) should to be linked together. Such a linking can be done manually by the researcher but we think that it is important to have an intuitive user interface to support the building of these *hypermodels*. To this end, in this paper we propose the paradigm of scientific workflows as the one to follow.

II. Scientific workflows

According to the Workflow Management Coalition (WfMC), workflows can be defined as "the automation of a business process, in whole or part, during which documents, information or tasks are passed from one participant to another for action, according to a set of procedural rules" [3]. This definition perfectly applies to workflows for the business community. More generally a workflow can be described as a sequence of operations or tasks needed to manage a business process or a computational activity. The latter definition can also be applied to scientific workflows, which are meant to decompose complex scientific experiments into a series of repetitive computational steps that could be run on supercomputers or distributed on a cloud system [4]. In our domain of interest these computational tasks correspond to in silico models and other computational tools that are executed in pipelines (or more generally in graphs) and communicate by exchanging data and parameter values.

There are already some popular workflow management systems, especially for bioinformatics. Taverna [5] is probably the most well-known one and recently has augmented its desktop version with a social networking website where the users can share their Taverna based workflows [6]. A complete web based workflow management system is Galaxy that features a user friendly, intuitive, "drag-and-drop" workflow editing functionality [7]. In this paper we describe a prototype for the TUMOR workflow editing and executing environment. This work revisits and extends similar efforts [8] that we made in past projects, focusing more on the modern web technologies and the requirements and the challenges of the VPH research. The decision of building a new workflow environment instead of reusing any of the existing ones was made after the evaluation of the available workflow

* This work was supported in part by the EC project TUMOR (FP7-ICT-2009.5.4-247754)

The authors ({ssfak; sakkalis; kmarias}@ics.forth.gr) are with the Computational Medicine Laboratory, Institute of Computer Science, FORTH, Vassilika Vouton, Heraklion, Greece (Phone: +30 2810 391672 Fax: +30 2810 391428)

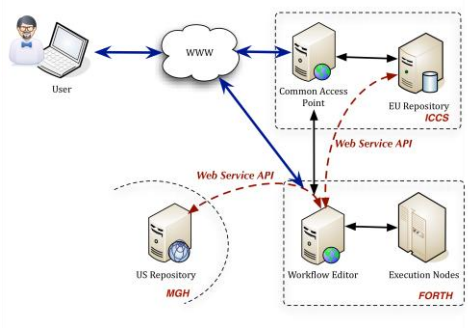


Figure 1 The TUMOR deployment architecture

management systems and the study of the TUMOR requirements. The most important factor for this decision was the particular architectural considerations and more specifically the requirements for integrating different model repositories with dynamic content that is frequently updated and different access restrictions. The adoption of a full web based deployment approach was also very important for this decision.

III. Architecture

The TUMOR environment is built as an online platform where its services and components are accessible over the World Wide Web [9]. The architecture therefore is designed with "service orientation" in mind, i.e. the software components expose a Web Service programmatic interface [10].

As a technical platform, the TUMOR environment consists of the following components (Fig. 1):

- The European Model and Data Repository: This is the "main" model repository, located in Europe. In addition to storing the cancer models of the European users and their anonymized data, this repository also maintains the users profile information.
- The US Model Repository: This is the American model repository, located in the US and operated by CViT. This is where US-CViT users store their models and data. It can be accessed from the European side but only the models can be transferred, due to the legal and ethical requirements.
- The Workflow Editing and Enactment environment, which is the web based application that allows the construction of simulation experiments through the linking of the available cancer models. In order to do so, the Workflow Environment accesses the EU and US model repositories and selectively retrieves models from their contents. It is hosted inside the EU and therefore it has access to the data stored in the EU repository. Nevertheless since it is a web application, it has to make authorization decisions based on the users profile in order to restrict the data access mechanisms only to the European users.
- The Common Access Point (CAP, for short): This is the main "entrance" to the platform. It is a web portal for interacting with the majority of the TUMOR services. Behind this portal there will be the EU Model and Data repositories and also the users profile database.

The model repositories are the primary sources for computational tasks in the workflow environment. The models are described in TumorML [11], an open markup language that provides descriptions of the input parameters and data of the models along side with curation (e.g. the creator of the model and his/her organization), classification (e.g. type of cancer), and other metadata.

IV. The TUMOR workflow environment

The workflow environment consists of two main components:

- The **workflow editor** (or designer) is a web application, accessible through the users' web browser. This is the graphical front-end for the editing of the workflows, the invocation of their execution, and the visualization of the results. A depiction of its interface can be found in Fig. 2.
- The **workflow engine** is the server side, which is responsible for the management and the execution of the workflows, the communication with the model repositories, etc.

This setting presents the workflow environment as a typical "Software as a Service" [12].

A. *Functionality of the Workflow Editor*

The workflow editor is the visual front-end to the application. It provides a "faceted" navigation and search [13] in the contents of the model repositories using the TumorML metadata. The following are the TumorML compliant criteria for filtering the models:

- Cancer type, e.g. Glioblastoma, lung, or breast. This allows to find models that are relevant only for the specific type of cancer.
- Mathematics type, e.g. whether discrete or continuous mathematic formulations are used for the implementation of the models
- Bio-complexity direction, to describe the macroscopic ("top down"), mesoscopic ("middle out"), microscopic ("bottom up") models
- Materialization model, i.e whether the tumor is assumed to be solid or liquid
- Homogeneity type, which accounts for the homogeneous or heterogeneous status of the simulated tumor.

Using these criteria the user is able to filter and select the ones that are more interesting, while free text search on models name and description is also supported. After

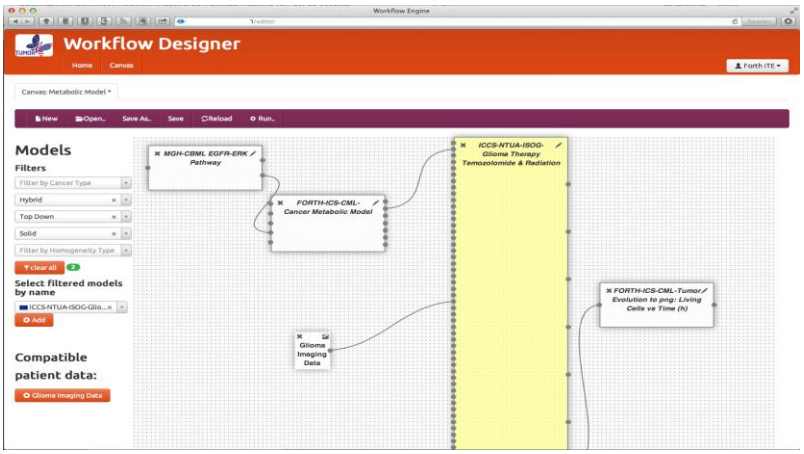


Figure 2. The main drawing area of the TUMOR workflow designer

selecting a model, the user can be included in a new or existing workflow and a visual representation of it appears in the central drawing area (the "canvas", as shown in Fig. 2). The user is then able to get a lot more information about the model, originating from its TumorML description. This information includes a free text description of the model, and the information about its required or optional input parameters and its outputs.

Easy and intuitive construction of the workflow is supported, by allowing the user to connect the models by dragging and connecting their inputs and outputs. Each input parameter (or possibly an output/result of a model) has a value from a restricted set of atomic datatypes, such as "integer", "float", and "string". The workflow editor makes sure that only compliant parameters are connected together and in fact it helps the user by providing a visual hint for the "matching" parameters when a user "drags" a connection. Another case where the workflow editor provides data type based matching information is when the user selects a model in the workflow drawing area. When this happens, the workflow backend checks whether the model requires some patient (anonymized) data and provides a suggestion for the user to include the matching data sets in the workflow and "connect" them with the model. The information for the matching datasets is of course provided by the models repository.

The user has always an immediate view of the model "readiness" to be executed. When the user selects "Run" from the menu he/she is presented with a form where all the parameters from all the models in the workflow are shown. The user can then provide explicit values for the parameters lacking one and after that trigger the execution. The execution of course happens in the background in the server side but the execution status, i.e. what models are currently running, is always available: the currently running models are shown "blinking" so that the user has immediate feedback on the execution progress. When the workflow has finished the user is presented with the results window where the results are shown inline is this can be the case (e.g. in the case of images, or plain numeric values). In addition to the final results of the workflow the user is able to see the immediate ones i.e. the output of the models that were in the start or the middle of the execution pipeline.

With respect to the implementation, the editor is using a lot of Javascript, and specialized Javascript libraries such as Jquery. Additionally a good use of of HTML5 [14] technologies is in place:

- Web Sockets [15] are used to support real time communications between the server and the client side. An example of such communication is during the monitoring of the execution status of a workflow. This almost realtime communication is delivered through the "web socket" connection so that any HTTP induced delays are eliminated.
- The WebStorage API [16] is used for the persistent data storage of key-value pair data. This is used for storing the local "session" of the user, so that refreshes to the drawing area or switching back and forth in other pages does not lose the content of the page. This can also be used to support undo/redo actions irrespective of the user leaving the page and coming back later. It is important to note that the information is stored locally - not in the server side - so saving and loading the "state" is quick and does not impose any burden (storage-wise) in the server. One of the disadvantages of this approach is that there are limitations on the size of the web storage imposed by the browsers (normally, around 5MBs).

B. Functionality of the Workflow Engine

The workflow engine is the server side of the environment and its main responsibilities are:

- The authentication of the users. This is subsequently delegated in the model repositories using the oAuth protocol[1], an open web based standard for authentication and authorization.
- The communication with the model repositories to retrieve the TumorML descriptions of the models

[1] http://oauth.org

and the corresponding executables and other data needed. This is implemented using SOAP based web services and REST [17].

- The storage and retrieval of user workflows. The persistence of the workflows is supported by a MongoDB database server[2].
- The execution of the workflows. The TumorML descriptions retrieved from the model repositories provide detailed information about the inputs and the outputs of each model. Using this information the workflow engine is the "orchestrator" of the models executions, deciding what to run next, and how to pass the data from the one model to the next.

The server side is implemented in Node.js, which is Javascript framework for networking applications based on the V8 javascript engine used in the Google Chrome browser and Chromium, its open source version [18]. Node.js is event based and (by default) single threaded but it is highly praised for its scalability for IO bound applications, e.g. network proxies and the majority of the web applications. We believe that the Workflow Engine is a perfect example for this type of applications since the main computationally heavy tasks are the models themselves at runtime. In essence, the models are executed in separate processes, as standalone command line executables, so this has no impact on the main workflow engine process. Of course the models should have been implemented like this, i.e standalone executables, which somehow restricts the model implementers. But in fact the approach is general because in TumorML there is information about how to get a whole "package" of the model that contains the required executables, libraries etc. and the command that the workflow engine or a human user needs to run in order to execute the model. In this way even Matlab scripts can be used as implementations assuming that there's matlab installation on the workflow server side and the executable is a wrapper script around the invocation to the matlab runtime.

V. Conclusions

Scientific workflows are important technical infrastructure for the implementation of in silico modelling in cancer research. In this paper we have described the design and a prototype implementation of a scientific workflow management system to support research in this area. The final system will need of course to be evaluated and validated by the research community but in any case it will distributed as open source software and provided as open access service over the web.

References

[1] D. Hanahan and R. Weinberg, "The hallmarks of cancer," Cell, vol. 100, no. 1, pp. 57 –70, 2000, ISSN: 0092-8674.

[2] T. Deisboeck and G. Stamatakos, Multiscale cancer modeling. CRC Press, 2010, vol. 34.

[3] D. Hollingsworth and U. Hampshire, "Workflow management coalition the workflow reference model," 1993.

[4] A. Belloum, E. Deelman, and Z. Zhao, "Scientific workflows," Scientific Programming, p. 171, 2006.

[5] T. Oinn, M. Addis, J. Ferris, D. Marvin, M. Senger, M. Greenwood, T. Carver, K. Glover, M. Pocock, A. Wipat, et al., "Taverna: a tool for the composition and enactment of bioinformatics workflows," Bioinformatics, vol. 20, no. 17, pp. 3045–3054, 2004.

[6] C. Goble and D. De Roure, "myExperiment: social networking for workflow-using e-scientists," in Pro- ceedings of the 2nd workshop on Workflows in support of large-scale science, ACM, 2007, pp. 1–2.

[7] J. Goecks, A. Nekrutenko, J. Taylor, et al., "Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences," Genome Biol, vol. 11, no. 8, R86, 2010.

[8] S. Sfakianakis, L. Koumakis, G. Zacharioudakis, and M. Tsiknakis, "Web-based authoring and secure enactment of bioinformatics workflows," in Grid and Pervasive Computing Conference, 2009. GPC'09. Work- shops at the, IEEE, 2009, pp. 88–95.

[9] S. Sfakianakis, V. Sakkalis, K. Marias, G. Stamatakos, S. McKeever, T. Deisboeck, and N. Graf, "An architecture for integrating cancer model repositories," in Conf Proc IEEE Eng Med Biol Soc. 2012, IEEE, 2012, pp. 6828–31.

[10] F. Curbera, M. Duftler, R. Khalaf, W. Nagy, N. Mukhi, and S. Weerawarana, "Unraveling the web services web: an introduction to SOAP, WSDL, and UDDI," Internet Computing, IEEE, vol. 6, no. 2, pp. 86–93, 2002.

[11] D. Johnson, J. Cooper, and S. McKeever, "TumorML: concept and requirements of an in silico cancer modelling markup language," in Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, IEEE, 2011, pp. 441– 444.

[12] P. Buxmann, T. Hess, and S. Lehmann, "Software as a service," Wirtschaftsinformatik, vol. 50, no. 6, pp. 500–503, 2008.

[13] D. Tunkelang, "Faceted search," Synthesis Lectures on Information Concepts, Retrieval, and Services, vol. 1, no. 1, pp. 1–80, 2009.

[14] D. Hyatt and I. Hickson, "HTML 5," W3C, W3C Working Draft, Mar. 2012. [Online]. Available: http://www.w3.org/TR/html5/.

[15] I. Hickson, "The web socket API," W3C, W3C Candidate Recommendation, Sep. 2012. [Online]. Available: http://www.w3.org/TR/websockets/.

[16] I. Hickson, "Web storage," W3C, Candidate Recommendation, Dec. 2011. [Online]. Available: http://www. w3.org/TR/webstorage/.

[17] R. Fielding and R. Taylor, "Principled design of the modern web architecture," ACM Transactions on Internet Technology (TOIT), vol. 2, no. 2, pp. 115–150, 2002.

[18] S. Tilkov and S. Vinoski, "Node. js: using Javascript to build high-performance network programs," Internet Computing, IEEE, vol. 14, no. 6, pp. 80–83, 2010.

[2] http://www.mongodb.org

APPENDICES

ANNEX

ONGOING RESEARCH

3D High Throughput in vitro Cancer Models

Feng Xu, Yasemin Akay and Metin Akay, *Senior Member, IEEE*

***Abstract*—Cancer is the second leading cause of death worldwide, with an estimated 1,437,180 new cases and 565,650 deaths in 2008 in US [1]. The development of effective therapeutics requires the understanding of the cellular mechanism and evolution of cancer. Although several research groups have developed 2D and 3D cancer models, these models have showed limited success in treatment response [2-5]. In addition to these models, we recently proposed a novel 3D platform to study the cellular mechanism at single cell level. We are currently investigating the potential use of it to study the cellular behavior of cancer cells to better understand the development of tumor evolution and metastasis and structural control, as well as external factors' influence on tumors.**

I. Introduction

Previous studies showed that altered microenvironments in three-dimensional (3D) model such as cell–extracellular matrix interactions may enhance tumor aggressiveness [6,7]. However, two-dimensional (2D) cancer models are not capable of representing these conditions [8-10]. Furthermore, cancer cell gene expression profiles [5] and response to mitogenic factors [11] have shown significant differences in 2D and 3D models. The complexity of 3D models has been uggested to be close to the in vivo situation compared to 2D models [12]. The 3D cancer models have also been shown to be effective in recapitulating some conditions encountered by tumor cells in vivo, which may not be case in standard 2D culture conditions [13,14,15].

Furthermore, several recent studies suggested that 3D cellular constructs can offer a more in-vivo like microenvironment than 2D monolayer cell culture. The biocompatible scaffold provides a similar architecture that resembles natural extracellular matrix (ECM). This enables a better free diffusion of various materials secreted by cells such as nutrients, enzymes, and growth factors. Besides, cells can also respond to the physical topography and stiffness of ECM, therefore, resulting in a more native cellular behavior as cell adhesion, migration, and proliferation like that within living states. For instance, Martin et al. [16] showed that 3D prognostic breast cancer signature can accurately predict clinical outcome compared 2D cases across independent datasets. Therefore, the ability to in vitro culture cells in 3D geometry allows investigation of cellular behavior in a more physiologically relevant state.

Although various 3D cancer models have developed, there are several challenges associated with these models. For instance, it is challenging for these methods to control the microarchitecture of the cellular construct model while cells are in a specific spatial distribution with microscale resolution in vivo. Besides, the existing methods have limited throughput, while large number of cancer constructs are needed for their applications (e.g., screening) to get statistical and meaningful results. To overcome all these challenges, there is a growing need to develop innovative high throughput 3D in vitro cancer models, which can regenerate distinct niches under well-defined conditions in a reproducible manner.

II. In vitro Single Cell Approach

Single cell analysis finds widespread applications in cellular development, genomics/proteomics analysis, and environment-interactive signaling and so on. Cell investigations conducted with large populations of cells would only reflect average values, but not consider the nature of cellular heterogeneity.

In vitro single cell analysis techniques are emerging as a powerful method to unravel cellular complexity with consideration to cellular heterogeneity. With advances in micro- and nano-techniques, many methods (e.g., microfluidics [17], surface modification [18], bioprinting [19], dielectrophoresis [20] and optical tweezers [21]) have been developed to unravel various cellular processes, such as cell differentiation [22], tumor progression and cell- cell/environment interaction [23], at the single cell level. However, these methods are currently based on 2D monolayer cell culture, which may not recapitulate the native situation. Therefore, it is of great importance to develop an effective method for single cell capture and analysis in a 3D microenvironment.

In our previous study, we developed a simple two-step lithography method to generate single cell captured in a 3D photo-polymerizable GelMA ring at high throughput [26]. Single cell capture with precise spatial control and high efficiency (~ 46.4%) was achieved by adjusting gap size and cell concentration. We applied this method to capture and culture single neurons, and found that neural axons grew out in 3D and formed axonal circles. These results indicate that that our method could be an enabling tool to analyze axonal development and autapse formation (an unusual type of synapse generated by a neuron on itself). In addition, the platform allows for the isolation of the axons from the neural soma, as well as the

FX is with the Biomedical Engineering and Biomechanics Center, The Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an 71049, P.R. China

YA, AD, and MA are with the Department of Biomedical Engineering, Cullen College of Engineering, Houston, TX 77204 USA

All correspondences address to Metin Akay at makay@uh.edu

investigation of axon development at single neuron level, which may provide fundamental information relating to many cellular processes, such as axonal development, synaptic plasticity and neural signal transmission.

It is known that tumor cells' genomes quickly become twisted in unusual ways as tumor cells evolve. These previous studies suggested the need for a feasible method to study cellular behavior of cancer cells (e.g., cancer heterogeneity) in a physiologically relevant state. Our method is capable of capturing and culturing single cancer cells, which could enable us investigate the tumor-genesis at individual cell scale, monitor the development of tumor evolution and metastasis, and structural control as well as external factors' influence on tumors. Recent advances in hydrogel chemistry and cell biology would enable us to incorporate molecules (such as chemoattractants or growth factors) to hydrogels to modulate cellular behaviors [27]. GleMA hydrogel possesses a large percentage of functional groups (such as carboxylic, amino and hydroxyl groups), which allows GelMA to be covalently modified or non-covalently mixed with growth factors. The trapping pattern of our method can also be easily modified by designing different photomasks to capture multiple number and types of cells in designed capture dots of a GelMA ring. This could be an enabling tool for researchers to study temporally controlled cell-cell interaction, e.g., cancer cell- healthy cell interaction to mimic tumor invasion.

III. Conclusion

We believe that our 3D high throughput neuron platform can also be used for other cellular mechanisms. We are currently investigating the potential use of it to study the cellular behavior of cancer cells to better understand the development of tumor evolution and metastasis and structural control, as well as external factors' influence on tumors.

References

[1] Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. CA: a cancer journal for clinicians 2008;58:71-96.

[2] Sasser AK, Sullivan NJ, Studebaker AW, Hendey LF, Axel AE, Hall BM. Interleukin-6 is a potent growth factor for ER-{alpha}-positive human breast cancer. FASEB J 2007;21:3763-70.

[3] Fiebig HH, Maier A, Burger AM. Clonogenic assay with established human tumour xenografts: correlation of in vitro to in vivo activity as a basis for anticancer drug discovery. Eur J Cancer 2004;40:802-20.

[4] Cukierman E, Pankov R, Stevens DR, Yamada KM. Taking cell-matrix adhesions to the third dimension. Science (New York, NY 2001;294:1708-12.

[5] Ghosh S, Spagnoli GC, Martin I, et al. Three-dimensional culture of melanoma cells profoundly affects gene expression profile: a high density oligonucleotide array study. Journal of cellular physiology 2005;204:522-31.

[6] Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000;100:57-70.

[7] Harris AL. Hypoxia--a key regulatory factor in tumour growth. Nature reviews 2002;2:38-47.

[8] Debnath J, Brugge JS. Modelling glandular epithelial cancers in three-dimensional cultures. Nature reviews 2005;5:675-88.

[9] Fischbach C, Chen R, Matsumoto T, et al. Engineering tumors with 3D scaffolds. Nature methods 2007;4:855-60.

[10] Pickl M, Ries CH. Comparison of 3D and 2D tumor models reveals enhanced HER2 activation in 3D associated with an increased response to trastuzumab. Oncogene 2008.

[11] Miralem T, Steinberg R, Price D, Avraham H. VEGF(165) requires extracellular matrix components to induce mitogenic effects and migratory response in breast cancer cells. Oncogene 2001;20:5511-24.

[12] Sutherland RM. Cell and environment interactions in tumor microregions: the multicell spheroid model. Science (New York, NY 1988;240:177-84.

[13] Phillips TM, McBride WH, Pajonk F. The response of CD24(-/low)/CD44+ breast cancer-initiating cells to radiation. Journal of the National Cancer Institute 2006;98:1777-85.

[14] Zhou J, Zhang H, Gu P, Bai J, Margolick JB, Zhang Y. NF-kappaB pathway inhibitors preferentially inhibit breast cancer stem-like cells. Breast cancer research and treatment 2008;111:419-27.

[15] Karnoub AE, Dash AB, Vo AP, et al. Mesenchymal stem cells within tumour stroma promote breast cancer metastasis. Nature 2007;449:557-63.

[16] Martin KJ, Patrick DR, Bissell MJ, Fournier MV. Prognostic breast cancer signature identified from 3D culture model accurately predicts clinical outcome across independent datasets. PLoS ONE 2008;3:e2994.

[17] Gao J, Yin XF, Fang ZL. Integration of single cell injection, cell lysis, separation and detection of intracellular constituents on a microfluidic chip. Lab Chip 2004;4:47-52.

[18] Palyvoda O, Bordenyuk AN, Yatawara AK, et al. Molecular organization in SAMs used for neuronal cell growth. Langmuir 2008;24:4097-106.

[19] Liberski AR, Delaney JT, Schubert US. "One cell-one well": a new approach to inkjet printing single cell microarrays. ACS Comb Sci;13:190-5.

[20] Jiang D, Sims CE, Allbritton NL. Microelectrophoresis platform for fast serial analysis of single cells. Electrophoresis;31:2558-65.

[21] Schutze K, Lahr G. Identification of expressed genes by laser-mediated manipulation of single cells. Nat Biotechnol 1998;16:737-42.

[22] Franco CB, Chen CC, Drukker M, Weissman IL, Galli SJ. Distinguishing mast cell and granulocyte differentiation at the single-cell level. Cell Stem Cell;6:361-8.

[23] Navin N, Kendall J, Troge J, et al. Tumour evolution inferred by single-cell sequencing. Nature;472:90-4.

[24] Hui EE, Bhatia SN. Micromechanical control of cell-cell interactions. Proc Natl Acad Sci U S A 2007;104:5722-6.

[25] Hutmacher DW. Biomaterials offer cancer research the third dimension. Nat Mater;9:90-3.

[26] Fan Y, Xu F, Huang G, Lu TJ, Xing W. Single neuron capture and axonal development in three-dimensional microscale hydrogels. Lab on a chip 2012.

[27] Wylie RG, Ahsan S, Aizawa Y, Maxwell KL, Morshead CM, Shoichet MS. Spatially controlled simultaneous patterning of multiple growth factors in three-dimensional hydrogels. Nat Mater 2011;10:799-806.

AUTHOR INDEX
(Open Access Version)