Proceedings
of the
4th International Advanced Research Workshop on In Silico Oncology and Cancer Investigation – The ContraCancrum Workshop

Athens, Greece, 8-9 September 2010

Edited by
Georgios Stamatakos and Dimitra Dionysiou

Available at www.4th-iarwisoci.iccs.ntua.gr

European Commission Information Society and Media

ContraCancrum

EMB IEEE

IFMBE

This volume is dedicated to cancer patients

G. Stamatakos and D. Dionysiou (Eds): *Proc. 4th Int. Adv. Res. Workshop on In Silico Oncology and Cancer Investigation (4th IARWISOCI) – The ContraCancrum Workshop*, Athens, Greece, Sept. 8-9, 2010 (www.4th-iarwisoci.iccs.ntua.gr)

PREFACE

Cancer is a *natural phenomenon* and as such it should be amenable to mathematical and computational description. Clinically driven complex multiscale cancer models can produce rather realistic spatio-temporal simulations of concrete clinical interventions such as radio-chemotherapy applied to individual patients. 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 enhance individualized treatment optimization. The latter constitutes 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 insight into tumor dynamics and optimization of clinical trial design and interpretation constitute short- and mid-term goals of this new domain.

The 4th International Advanced Research Workshop on *In Silico* Oncology and Cancer Investigation (IARWISOCI) that was held in Athens, Greece on September 8-9, 2010 proved an excellent opportunity for both shaping and advancing the discipline. The event, being also the ContraCancrum project workshop, attracted leading researchers from around the globe and from a multitude of fields involved in *in silico* oncology. The present volume includes their contributions in the form of short papers.

The workshop was sponsored by the European Commission, Directorate-General for Information Society and Media - Virtual Physiological Human (VPH) initiative, through the ContraCancrum project (www.contracancrum.eu). It was also technically co-sponsored by the Institute of Electrical and Electronics Engineers (IEEE), Engineering in Medicine and Biology Society (EMBS) (http://www.ieee.org/conferences_events/conferences/conferencedetails/index.html?Conf_ID=17234) and endorsed by the International Federation for Medical and Biological Engineering (IFMBE) (http://ipem.ac.uk/ifmbe_ngen_public/default.asp?ID=929).

Each submitted technical program paper underwent a review by three peers (two reviewers and one review coordinator). In order for the whole process to meet the IEEE-EMBS reviewing standards a memorandum of understanding was signed by the General Chair of the workshop and IEEE-EMBS in early 2010. The latter contains inter alia all the IEEE approved reviewing process details.

I would like to thank the Members of the Organizing Committee for their crucial contribution to the success of the event, the reviewers of the submitted manuscripts for their valuable feedback, my colleague Dr Dionysiou for her important contribution to the editorial process and all contributing authors and workshop speakers for their excellent work, inspiring presentations and fruitful discussions. Special thanks are due to the European Commission, the ContraCancrum project, IEEE-EMBS and IFMBE for the technical and/or financial (co-) sponsorship of the event. The enthusiasm of the involved administrative staff is duly acknowledged.

I hope that the present volume will turn out to be a useful tool for both the advancement of *in silico* oncology and the achievement of its mid- and long-term *patient centered* goals.

Athens, 17 November 2010

Georgios S. Stamatakos
General Chair, 4th IARWISOCI
Research Professor
Institute of Communication and Computer Systems-
National Technical University of Athens, Athens, Greece

G. Stamatakos and D. Dionysiou (Eds): *Proc. 4th Int. Adv. Res. Workshop on In Silico Oncology and Cancer Investigation (4th IARWISOCI) – The ContraCancrum Workshop*, Athens, Greece, Sept. 8-9, 2010 (www.4th-iarwisoci.iccs.ntua.gr)

PAPER CITATION FORMAT

(Co-)Author(s), "Paper Title." In G. Stamatakos and D. Dionysiou (Eds): *Proc. 4th Int. Adv. Res. Workshop on In Silico Oncology and Cancer Investigation (4th IARWISOCI) –The ContraCancrum Workshop*, Athens, Greece, Sept. 8-9, 2010 (www.4th-iarwisoci.iccs.ntua.gr), pp.xx-xx.

It is noted that the Proceedings are freely available on the workshop website www.4th-iarwisoci.iccs.ntua.gr

CORRESPONDENCE

All corresponcence should be addressed to

Dr Georgios S. Stamatakos
In Silico Oncology Group
Laboratory of Microwaves and Fibre Optics
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

CONTENTS

VII. Computing and Multiparametric Exploration Technologies

VIII. Image Processing for Tumor Modeling

IX. Multiscale Data and Simulation Tool Handling Technologies

X. Markup Languages and Collaborative Systems for Cancer Model Development

SHORT INTRODUCTORY PAPERS

G. Stamatakos and D. Dionysiou (Eds): *Proc. 4th Int. Adv. Res. Workshop on In Silico Oncology and Cancer Investigation (4th IARWISOCI) – The ContraCancrum Workshop*, Athens, Greece, Sept. 8-9, 2010 (www.4th-iarwisoci.iccs.ntua.gr)

In Silico Oncology and the Oncosimulator Concept: A Platonic Approach to Medical Science

Georgios S. Stamatakos

***Abstract*— A brief introduction to the emergent discipline of *in silico* oncology along with its predominantly Platonic character is provided. The conceptual origins of the *oncosimulator*, one of the key notions and constructs of the discipline are traced back to antiquity.**

I. Introduction

In silico oncology is a complex and multiscale combination of sciences and technologies intending to simulate malignant tumor growth and tumor and normal tissue response to therapeutic modalities at all scales of biocomplexity. The long term goal of *in silico* oncology is to quantitatively understand cancer and related phenomena and optimize therapeutic interventions by performing *in silico* (i.e. on the computer) experiments based on the individual patient's clinical, imaging, histopathological, molecular and pharmacogenomic data [1]. In order to achieve such an ambitious goal translation of cancer models into the clinical trials arena is a *sine qua non* condition.

II. The Platonic Character of *In Silico* Oncology

The central axiom of *in silico* oncology is that *cancer is a natural phenomenon and as such it should be amenable to mathematical and computational description*. But a mathematical and computational description presupposes abstraction. Therefore, in order to end up with a domain of practical value the following chain of processes has to be implemented:

Reality → Abstraction → Virtuality → Virtual Processing (Virtual Patient Treatment) → Treatment Optimization → Reality → Actual Patient Treatment → Feedback to Virtuality → Optimized Virtuality (Optimized Treatment Simulator)

The underlined entities and processes can be seen as essentially belonging to the Platonic world of ideas or the idealistic domain. However, since reality based virtuality has to always be strictly validated against reality it is obvious that a combination of both a Platonic and an Aristotelian (i.e. reality focused) approach rather than a pure Platonic perspective is necessary for a trustable and useful outcome.

III. The Oncosimulator and Its Conceptual Precursors

The *oncosimulator* is at the same time a concept of multiscale integrative cancer and treatment affected normal tissue biology, an algorithmic construct and a software tool which aims at supporting the clinician in the process of optimizing cancer treatment on the patient individualized basis. Additionally it is a platform for better understanding and exploring the natural phenomenon of cancer as well as training doctors and interested patients alike. In order to achieve all of these goals it must obviously undergo a thorough clinical optimization and validation process.

The idea of simulating natural phenomena can be traced back to antiquity. The Antikythera mechanism [2-3], an ancient mechanical computer designed to simulate planetary motions and calculate astronomical positions, has been widely seen as the first technological simulator. It is thought to have been built about 150–100 BC. Simulation has nowadays become a central element of every conceivable scientific domain, medicine being no exception. The development of oncosimulators is a testimony to this. However, the predictive power of oncosimulators remains to be clinically validated. This is one of the fundamental tasks of *in silico* oncology research being currently in progress.

References

[1] G. Stamatakos , "Spotlight on Cancer Informatics," Cancer Informatics, 2 , 83-86, 2006. http://www.la-press.com/article.php?article_id=120
[2] http://en.wikipedia.org/wiki/Antikythera_mechanism
[3] The Antikythera Mechanism Research Project. http://www.antikythera-mechanism.gr/

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

Scientific and Technological Background of *In Silico* Oncology

Nikolaos Uzunoglu

***Abstract*—The analysis of highly complex phenomena such as cancer requires a multidisciplinary approach relying on a successful combination and cross-polination of several scientific and technological domains. This short paper briefly outlines the recruitment of several scientific and technological fields such as theoretical physics and numerical techniques in order to support the development of *in silico* oncology in the Laboratory of Microwaves and Fiber Optics of the National Technical University of Athens. A short outline of a modeling approach to tumor response to radiation therapy is provided as an multidisciplinary endeavor paradigm.**

I. Multidisciplinarity in Science

The scientific discipline of electrical engineering arose about 120 years ago just after the emergence of the first applications of Maxwell's electromagnetic theory. The foundations of electromagnetic theory are very similar to the Newtonian theory of dynamics and essentially both physical theories have a structure very similar to Euclidean geometry. The beauty of the electromagnetic theory has been enhanced by its wide applicability in explaining the dynamical behavior of all phenomena related to electric and magnetic fields varying along the time axis. Furthermore, when the Maxwell electromagnetic theory was recruited in order to explain the interaction of matter with electromagnetic energy its failure to explain the involved phenomena led to the emergence of quantum theory while electromagnetism in moving media gave the way to the special theory of relativity and afterwards to the theory of general relativity.

The success of Maxwell theory is based on its reductionist concepts of scientific thought. What we call science today was born in western Anatolia and in particular in the town of Miletus from the 6th century B.C. onwards. It was based on the cultivation of the reductionist method. The observation of natural phenomena in combination with a pioneering spirit resulted into the concepts through which human mind can model nature by exploiting the power of thought. In this way by explaining simple phenomena one can build up a scientific method able to explain complex phenomena that may be viewed as combinations of the simpler ones.

The research activities in the framework of electrical engineering science (and latter on also in computer science and informatics) has been much influenced by the previously mentioned scientific way of thinking. Especially during the 1960's many fundamental theories of physics such as applied quantum mechanics and quantum electrodynamics paved the way to many new inventions and technological applications. Advances in computer technology during the same period provided the opportunity to numerically model complex structures involving the interaction of matter with radiation. Problems considered impossible to analyze until that time became readily solvable and accurate predictions of the systems' behavior were obtained. In the mid-1990's it was possible to numerically compute electromagnetic structures with complexity of high degree. The main thrust to these efforts came from the calculation of the radar signals that were reflected from flying objects. This spirit of exploiting the capability of performing large scale computations penetrated the research community working on electromagnetics modeling. However, this advance did not come without any adverse effects since in many cases the physical intuition started to gradually disappear.

Meanwhile the field of medical physics that was initiated as early as 1950s started to expand into new domains of medical applications. The technological applications in medicine, being restricted in the 1940's to only X-ray screening and a few electrical measurements, was continuously expanding to new fields such as linear accelerator applications, computers in medicine and to an extensive degree computers in biological process modeling.

However, even by the end of 1990's it was difficult to propose a detailed multiscale model of malignant cell cluster behavior relying on computational methods for funding. In many cases proposals were rejected with the dogmatic statement "*cancer cannot be modeled by mathematics and physics*".

Research on biomedical engineering at the Microwave and Fiber Optics (MFOL)– National Technical University of Athens started in 1985 in two specific topics i.e hyperthermia and brain modeling. The influence of the above mentioned reductionist attitude combined with an engineering approach was implemented in order to design and construct new treatment systems. This led to the

N.K.Uzunoglu is with the School of Electrical and Computer Engineering, National Technical University of Athens, 9 Iroon Polytechniou, 157 80 Zografou, Greece (phone: + 30 210 772 3556, fax: + 30 210 772 3557, e-mail: nnap@otenet.gr)

G. Stamatakos and D. Dionysiou (Eds): *Proc. 4th Int. Adv. Res. Workshop on In Silico Oncology and Cancer Investigation (4th IARWISOCI) – The ContraCancrum Workshop*, Athens, Greece, Sept. 8-9, 2010 (www.4th-iarwisoci.iccs.ntua.gr)

combination of theoretical modeling work with immediate clinical applications in the two aforementioned fields.

Hyperthermia is an adjuvant cancer treatment modality based on the cytotoxic properties of biologically high temperatures (43 ^{o}C) achieved within cancer cells and tissues. Hyperthermia is usually achieved by using microwaves to heat the tumor. Soon it was realized that in order to achieve the optimal treatment it was necessary to model the electromagnetic and thermal response of the anatomic area of interest in conjunction with the response of the region of interest to radiation therapy. This led the MFOL research team to develop detailed computational models based on the past experience of electromagnetic computations. A similar approach was adopted for the case of brain studies since the problem to be dealt with in brain modeling was again of enormous complexity. Electromagnetics methods proved to be very useful and several brain response phenomena were succesfully treated.

Based on the above concepts and after having received stimuli by medical doctors and in particular by Prof. N. Zamboglou, Klinikum Offenbach, Germany as well as by the pioneering work of Prof.W. Duchting, University of Siegen, Germany, an effort started in the late 1990's with the aim to model in detail the dynamic response of cancerous cell clusters (initially *in vitro* and subsequently *in vivo*). Despite the initial negative precondition of numerous medical doctors the recognition of the fact that many practices in clinics fail because of the lack of accurate knowledge of systems biology helped in the acceptance of the proposed detailed methods in cancer modeling. After a decade a change of mood occurred and many former unbelievers of the potential of cancer modeling began to cooperate with engineering and physics groups so that several interdisciplinary research teams are presently intensively working on the topic. The long term goal is to develop individualized treatment methods taking into account all the biological details in each patient case. There is substantiated hope that this approach will provide concrete results in achieving optimization of the therapeutic methods in the clinical environment.

II. The Paradigm of Modeling the Response of Tumor to Radiation Therapy

In the radiation therapy setting current treatment planning algorithms are based on the concept of physical optimization of the dose distribution and rely on rather crude biological models of tumor and normal tissue response. Such algorithms practically ignore the highly complicated dynamic behavior of malignant and normal cells and tissues. The introduction of advanced biosimulation methods based on cell proliferation mechanisms and also on information drawn from the cellular and molecular properties of each individual malignancy and each individual patient are expected to substantially improve the radiation therapy efficiency. This would be accomplished by using alternative fractionations, spatial dose distributions and even combination with other therapeutic modalities such as chemotherapy, hyperthermia etc. Therefore, efficient modeling, simulation and visualization of the biological phenomena taking place before, during and after irradiation is of paramount importance. Discrete time algorithmic descriptions (simulations) of the various phenomena offer the possibility of taking into account a large number of involved mechanisms and interactions. The same philosophy has already been extensively applied to purely technological problems and the emerged numerical methods (e.g. the Finite Difference Time Domain (FDTD) technique) have proved to be very efficient and reliable. A further prominent characteristic of the biological phenomena under consideration is stochasticity. The fate of a single irradiated cell cannot be accurately predicted for example. Only survival probabilities can be assigned to the cell based on the accumulated experimental and clinical observations made on large cell populations. Furthermore, the exact spatiotemporal distribution of the various cell cycle phases within the tumor volume is generally unknown, although some plausible macroscopic hypotheses can be made. Therefore, stochastic techniques such as the generic Monte Carlo method seem to be particularly appropriate for the prediction of tumor growth and response to radiation therapy. The practical usefulness of such methods is both to improve understanding of the cancer behavior and to optimize the spatiotemporal treatment plan by performing *in silico* (= on the computer) experiments before the actual delivery of radiation to the patient. The clinician would be able to perform computer simulations of the likely tumor and adjacent normal tissue response to different irradiation scenarios based on the patient's individual imaging, histologic and genetic data. The simulation predictions would eventually support him or her in selecting the most appropriate fighting strategy. The basic approach is to develop detailed four dimensional simulation models of the biological systems under consideration whereas at the same time to make use of advanced technologies (e.g. visualization systems, parallelization etc.). A grossly similar approach has been implemented in the case of chemotherapy.

III. Conclusions

Indicative contributions of various scientific domains to the development of *in silico* oncology have been outlined. The importance of cross-pollination of scientific and technological ideas and approaches has been revealed through the paradigm of the MFOL's initiatives in *in silico* oncology. Finally, a high level categorization of simulation

techniques applicable to the paradigmal case of radiation therapy has been presented and discussed.

SHORT RESEARCH PAPERS

I. Clinical Aspects

G. Stamatakos and D. Dionysiou (Eds): *Proc. 4th Int. Adv. Res. Workshop on In Silico Oncology and Cancer Investigation (4th IARWISOCI) – The ContraCancrum Workshop*, Athens, Greece, Sept. 8-9, 2010 (www.4th-iarwisoci.iccs.ntua.gr)

Multiscale Clinical, Morphologic and Genomic Data Management for in Silico Oncology of NSCLC

Yoo-Jin Kim, Christian Veith, Jan Palm, Aleksandar Grgic, Robert Rixecker, Cristiana Roggia, Winfried Bauer, Günther Schneider, Dietmar Tscholl, Norbert Graf, and Rainer Maria Bohle

***Abstract*— Contra Cancrum is a multiscale platform that aims at developing a composite multilevel simulation model of malignant tumor development and tumor tissue response to therapeutic modalities and treatment schedules. The purpose is to better understand the *natural phenomenon* of cancer at all levels of biocomplexity and at the same time to support the *disease* treatment optimization procedure in the patient's individualized context. Non small cell lung cancer (NSCLC) is one of the diseases currently being analyzed in this platform. We describe the different ways of clinical guidance and collection of general patient data, morphologic (macropathologic, histopathologic, immunopathologic) and genomic data within the contra cancrum project.**

I. Introduction

The main goal of the *Contra Cancrum* platform is to build up a clinically proved data base for various cancer types. Currently a major focus is on NSCLC and glioblastoma. NSCLC is still one of the most aggressive cancer types of adults with a nearly unchanged very low 5-year survival rate (below 20%) since decades. To build up an in silico model for NSCLC offering opportunities to test real clinical situations, a virtual platform had to be fed – in a top-down approach - with most up to date clinical, morphologic and therapy-orientated genomic data. We show how this cancer type-specific data base was created, to serve as the basic platform for the evaluation of in silico and personalized approaches. All clinical, morphologic and genetic parameters used in this approach are presented and discussed.

Y. J. Kim is with the Institute of Pathology, Saarland University, Homburg/ Saar, Germany (e-mail: yoo-jin.kim@uks.eu),

C. Veith is with the Institute of Pathology, Saarland University, Homburg/ Saar, Germany (e-mail: christian.veith@uks.eu),

J. Palm is with the Department of Radiation Therapy and Radiation Oncology, Saarland University Hospital, Homburg/ Saar, Germany (e-mail: Jan.Palm@uks.eu),

A. Grgic is with the Department of Nuclear Medicine, Saarland University Hospital, Homburg/ Saar, Germany (e-mail: aleksandar.grgic@uks.eu),

R. Rixecker is with the Institute of Pathology, Saarland University, Homburg/ Saar, Germany (e-mail: robert@ludwigspark.de),

C. Roggia is with the Institute of Pathology, Saarland University, Homburg/ Saar, Germany (e-mail: roggia.christiana@uks.eu),

W. Bauer is with the Saarland Tumor Center, Saarland University Hospital, Homburg/ Saar, Germany (e-mail: winfried.bauer@ uks.eu),

G. Schneider is with the Department of Diagnostic and Interventional Radiology, Saarland University Hospital, Homburg/ Saar, Germany (e-mail: dr.guenther.schneider@uks.eu),

D. Tscholl is with the Department of Thoracic and Cardiovascular Surgery, Saarland University Hospital, Homburg/ Saar, Germany (e-mail: dietmar.tscholl@uks.eu),

N. Graf is with the Department of Pediatric Oncology and Hematology, Saarland University Hospital, Homburg/ Saar, Germany (e-mail: norbert.graf@uks.eu),

R. M. Bohle is with the Institute of Pathology and with the Saarland Tumor Center, Saarland University and Hospital, Homburg/ Saar, Germany (corresponding author to provide phone: +49-6841-16-23850; fax +49-6841-16-21802; e-mail: rainer.bohle@uks.eu).

II. Clinical Parameters

A. Tumor Classification

According to the current guidelines of the WHO (1) the histopathological subtyping of NSCLC was performed. As squamous cell carcinomas and adenocarcinomas amount to more than 80% of all NSCLC, these two subtypes were separately used. All other types of NSCLC, i.e. large cell carcinomas, adenosquamous carcinomas etc. were put into a third group, named "other carcinomas". The NSCLC were graded according to the WHO criteria. Beyond, the tumor site was recorded, the ICD-O-codes were determined and the procedures leading to histopathological tumor classification (biopsy or resection specimen) were recorded. The data were submitted to the Saarland Cancer Registry. Patients were informed on data recording and could refuse recording according to the law of the state of Saarland.

B. TNM Classification

In all NSCLC cases, TNM stage grouping was performed according to the seventh edition of the TNM classification of malignant tumors (2). For resection specimen the pT and pN staging was used. Stage grouping using the "M" descriptors and for "biopsy cases" was performed after clinical and radiological investigations were completed and allowed clinical decisions.

C. Clinical Imaging

Clinical appearances of the lung cancer diseases were documented using CT and/ or MRI imaging. From many patients additional PET CT data were available. The

imaging files were recorded in DICOM image sets. At least preoperative and/ or pre-biopsy findings were integrated into the data base. If available, postoperative or post-chemo-/ post-radiotherapy images were also stored as DICOM files.

D. Therapy

Data documenting the kind of therapy (thoracic surgery, adjuvant, neo-adjuvant or palliative radio- and/ or chemotherapy) were recorded through the clinical cancer registry of the university hospital of Saarland. Modalities and dates of therapies were integrated into the data bases.

E. Follow-up

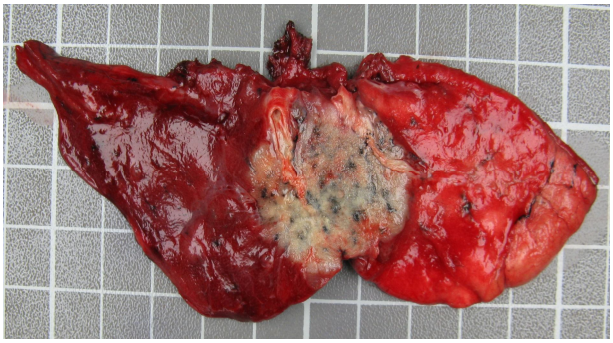
Clinical follow-up parameters were: tumor state (recurrence/ stable disease/ progress), rTNM or yTNM status, blood counts, occurrences and types of sequelae, location of recurrences or organ-metastasis in addition to second or third lines therapies.

III. Morphologic Parameters

A. Macroscopic Data

All resection specimens of lung cancers were analyzed after immediate transport to the Institute of Pathology. They were measured in three dimensions and dissected natively into 5 mm thick tissue slices. Tumor and tumor-free slices were photographed digitally (fig.1). Files were stored as tiff-files. An average number of four macrophotographs per resection specimen were recorded.

Fig. 1 : Macrophotograph of a resection specimen of NSCLC.



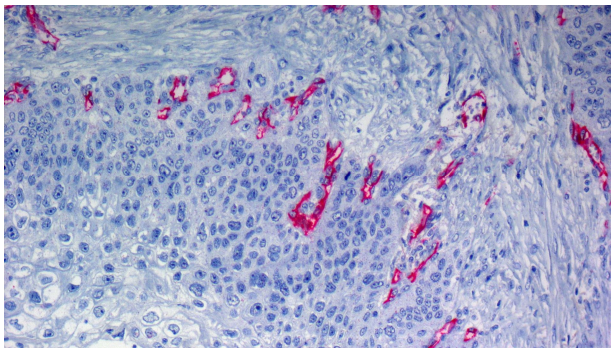
B. Microscopic Data

From all biopsy specimens and from resection specimens of NSCLC representative microphotographs were taken using a Zeiss photomicroscope together with the AxioVision imaging software (Zeiss, Göttingen, Germany). At least 5 microphotographs were taken from each case. Together with the data from *III A*, areas of tumor necrosis and/ or tumor regression could be determined.

C. Immunohistologic Data

Using the same equipment as for *III B*, up to five representative microphotographs were taken from Ki 67- (proliferation marker; clone MIB-1, Dako, Hamburg, Germany) and CD31- (angiogenesis marker; clone JC/70A, Dako) (fig. 2) immunostained tumor slides (3, 4).

Fig. 2 Microphotograph of an anti-CD31 immunostained NSCLC.

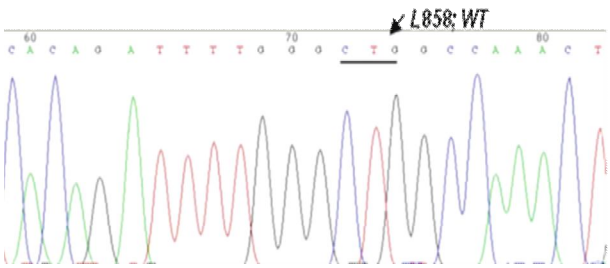


IV. Genomic Parameters

A. Epidermal Growth Factor Receptor (EGFR)

Exons 18, 19, 20, and 21 of the EGFR gene were analyzed from NSCLC tumor tissue (5) manually dissected or laser microdissected from 5 µm thick slides after microscopic control. Short amplicons (< 200bp) were generated after standard DNA extractions from formalin-fixed and paraffin embedded tissue and after PCR. Afterwards sequencing was performed using the Sanger sequencing technique (Seqlab, Göttingen, Germany). Sequencing data were analyzed using *Chromas Lite 2.01* software (freeware). Data recording focused on EGFR activating mutations and deletions or wild-type EGFR.

Fig. 3 DNA sequencing of NSCLC : electropherogram of an EGFR exon 21 amplicon.



B. K-RAS

K-RAS mutation analysis was done using DNA extraction from manually dissected or microdissected NSCLC tumor tissue and K-RAS exon 2 PCR (5). LCD-Array K-RAS 1.4 (Chipron, Berlin, Germany) served for specific detection of mutations in codon 12 and codon 13 of the KRAS gene. As all the mutations detected with this assay are KRAS-activating mutations, mutation status or wild-type status transferred into the data base.

V. DISCUSSION

The *Contra Cancrum* multiscale platform aims at developing a multilevel simulation model of cancer biology and in vivo tumor response to therapeutic modalities. From the scientific perspective the proposed platform provides a better understanding and exploration of tumor development and biology at different biocomplexity levels (6). The simulation models also aim to support the clinician in the process of optimizing cancer treatment and clinical decision making in individual patients (7).

The discoveries in molecular-genetic research offer a great opportunity in the development of an individualized cancer therapy. Computational models enable an integration of these molecular-genetic data into an amount of clinical and pathological data within complex scenarios of virtual trials. Nevertheless, a concept for an easy and anonymized access to these heterogeneous data is essential.

In our model, documentations of individual cancer therapy and follow-up were retrieved from the records of the clinical cancer registry of the university hospital of Saarland. This approach allows for a complete and standardized documentation and retrieval of clinical data.

The multi-scale approach of the *in silico* model captures tumor progression by taking into account ongoing molecular, cellular and tumor scale events. In the here proposed data management, the tumor scale is represented by the macromorphology (fig. 1) and the clinical imaging data. In addition to the microscopic morphology, cellular scale events are represented by tumor proliferation and angiogenesis, both of which have prognostic impact in NSCLC (4) and also contain information of growth patterns that result in different biological subtypes (3). We integrated these morphological parameters into the data base in terms of immunohistochemical expression levels of the proliferation-associated protein Ki 67 and in terms of microvessel density in CD31-immunostained tumor slides (fig. 2). On the level of molecular scale events, genomic parameters basically influence the tumor biology and also provide opportunities for individualized molecular-targeted therapies. There is mounting evidence that patient selection will play a key role in the successful development of any targeted agent. As recent studies have reported that clinical response to EGFR inhibitors is associated with activating mutations of EGFR in the advanced stage of lung cancer, we transferred EGFR sequencing data into our data base. As KRAS mutation was reported as a poor prognostic factor in the clinical outcome of early lung cancer after surgical resection (5), we also included KRAS-activating mutations, mutation status or wild-type status into the data set.

Until now, whether EGFR expression or absence of KRAS mutations are appropriate markers for targeted therapies is still under study. By applying *in silico* technologies, including virtual simulation of clinical trials faster estimations on treatment strategies can be achieved without performing time-consuming and expansive *in vivo* experimentations. The here proposed management of multiscale data exemplifies the assessment, retrieval and integration of heterogeneous and multilevel data (clinical, morphologic and genomic data as well as follow-up and treatment/ treatment response data) that are pivotal for establishing and performing an *in silico* oncology approach to NSCLC.

ACKNOWLEDGMENT

The project was partially funded by the European Commission, grant number: 223979. The authors thank Julia Michaely for her great technical help.

REFERENCES

[1] W. D. Travis, E. Brambilla, H. K. Müller-Hermelink, and C. C. Harris, Pathology and Genetics of Tumors of the Lung, Pleura, Thymus and Heart. WHO Classification of Tumors, Volume 10. 3rd ed., Lyon: IARC Press, 2004.

[2] L. H. Sobin, M. K. Gospodarowicz, and C. Wittekind, TNM Classification of Malignant Tumors. 7th ed., Oxford: Wiley-Blackwell, 2009.

[3] P. Sardari Nia, C. Colpaert, P. Vermeulen, J. Weyler, F. Pezzella, P. Van Schil, and E. Van Marck, "Different growth patterns of non-small cell lung cancer represent distinct biologic subtypes," *Ann. Thorac. Surg.*, vol. 85, no. 2, pp. 395-405, Feb. 2008

[4] T. C. Mineo, V. Ambrogi, A. Baldi, C. Rabitti, P. Bollero, B. Vincenzi, and G. Tonini, "Prognostic impact of VEGF, CD31, CD34, and CD105 expression and tumor vessel invasion after radical surgery for IB-IIA non-small cell lung cancer," *J. Clin. Pathol.* vol. 57, no. 6, pp. 591-597, Jun. 2004.

[5] Y.-T. Kim, T.-Y. Kim, D.-S. Lee, S.-J. Park, J.-Y. Park, S.-J. Seo, H.-S. Choi, H.-J. Kang, S. Hahn, C.-H. Kang, S.-W. Sung, and J.-H. Kim. "Molecular changes of epidermal growth factor receptor (EGFR) and KRAS and their impact on the clinical outcomes in surgically resected adenocarcinoma of the lung," *Lung Cancer*, vol. 59, no. 1, pp. 111-118, Jan. 2008.

[6] 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.

[7] Graf N, Hoppe A, Georgiadi E, Belleman R, Desmedt C, Dionysiou D, Erdt M, Jacques J, Kolokotroni E, Lunzer A, Tsiknakis M, Stamatakos G: "In silico oncology for clinical decision making in the context of nephroblastom," *Klin. Padiatr*. vol. 221, no. 3, pp. 141-149, May 2009.

Virtual Simulation (Preplanning) in Interventional Radiation Therapy. The Potential of *In Silico* Oncology (Abstract)

Nikolaos Zamboglou

***Abstract*— The objective of radiation treatment is to achieve the maximum dose within the tumor and at the same time the minimum dose to proximal normal tissues and organs. Improvements have been made in recent years through radiobiological studies which have shown that the effect of the radiation can be enhanced by high LET irradiation, by combined chemotherapy and radiotherapy, and by combined radiotherapy and hyperthermia.**

However, in clinical studies, as distinct from experimental radiobiology, the use of these radiation modifiers has not yet achieved their initial expectations and are therefore, not fully established within the armamentarium of radiation oncology. Currently the only certainty in radiotrherapy is that an increase in fractionation dose and in total dose can improve local control. Nevertheless, it must be emphasized that there can be a trade off in terms of incidence of complications when the dose levels are high.

Technical innovations, include three-dimensional conformal treatment planning, radiosurgery techniques, stereotactic irradiation, intraoperative teletherapy, and the use of imaging modalities such as CT and MRI. These all have improved the possibilities of external beam radiotherapy because they permit delivery of a higher dose to the target volume relative to the given to normal tissues and organs. Nevertheless, there are limitations in external beam radiotherapy, including the practically maximum volume which can be irradiated, the accuracy achievable when determining the target volume and the worldwide availability of linear accelerators. The latter is sometimes forgotten by large well established oncology centres, but the installation of linear accelerators worldwide is still far from optimal.

Interstitial brachytherapy has always been an alternative modality but until the 1990s there had been the following limitations which means that the technique has not been as widely applied as it deserves.

1. No widespread clinical experience in the applications of catheters for interstitial brachytherapy.

2. Prior to remote afterloading machines, the interstitial techniques of pre-loaded sources or manually afterloaded sources had presented major problems in radiation protection. The earliest remote afterloading machines were designed for LDR intracavitary applications, such as the Selectron-LDR, and for example the 137Cs pellet source size was too large for interstitial applications.

3. There had been insufficient documentation of doses to target volumes and tissues and organs at risk.

4. There had been no radiobiological approach which would permit the extrapolation of LDR interstitial experience to HDR interstitial techniques.

5. There had been time constraints for the geometrical construction of catheter configurations. It had not been unknown to take one person-day for such reconstructions when adequate computer software was not available. However, the limitations mentioned above can now largely be overcome. For example the use of MR imaging, ultrasound imaging, and particularly of CT imaging have opened new horizons for this important specialty of interstitial brachytherapy. Thus the placement of catheters can now be accurately and safely made in most organs of the body.

Within this framework a significant number of novel systems aiming at the imaging guided application of brachytherapy interstitial catheters within brain tumors, breast, soft tissue sarcomas, bone metastases etc. have been developed by our research team. To this end CT and/or MRI and/or PET and/or ultrasound based navigation systems have been developed. Isodose calculations are reliably performed based on extensive imaging information. Appropriate test phantoms have been devised and successfully applied to the planning validation process. Efficient fusion techniques (e.g. manual markers, contours, mutual information) have also been applied in order to maximize the imaging based usable information. Clinical experience so far has justified both the applicability and effectiveness of the previously mentioned systems.

However, although physical optimization of radiotherapy and brachytherapy has been significantly advanced over the past years, there is a great gap in the translation of the current biological knowledge into radiotherapeutic practice. Since the potential of the latter appears to be great we strongly support the development and future clinical translation of *in silico* oncology, a new domain focusing on multiscale cancer and affected normal tissue biology in the clinical context. Thus we expect that a combination of physical and multiscale biological optimization of radiotherapy and brachytherapy will considerably improve the therapeutic outcome.

N.C.Zamboglou is with the Radiation Therapy Clinic, Klinikum Offenbach, Starkenburgring 66, 63069, Offenbach, Germany (Tel. +49 6984053335; email: nikolaos.zamboglou@klinikum-offenbach.de, nzamboglou@klinikum-offenbach.de)

II. Biochemical and Molecular Level Based Modeling

G. Stamatakos and D. Dionysiou (Eds): *Proc. 4th Int. Adv. Res. Workshop on In Silico Oncology and Cancer Investigation (4th IARWISOCI) – The ContraCancrum Workshop*, Athens, Greece, Sept. 8-9, 2010 (www.4th-iarwisoci.iccs.ntua.gr)

Rapid and accurate ranking of binding affinities of epidermal growth factor receptor sequences with selected lung cancer drugs

Shunzhou Wan and Peter V. Coveney

***Abstract*—We investigate the binding affinities of two tyrosine kinase inhibitors – AEE788 and Gefitinib – to epidermal growth factor receptor (EGFR) using molecular dynamics simulation. The interactions between these inhibitors and the EGFR kinase domain are analyzed using multiple short (ensemble) simulations and the Molecular Mechanics/Poisson-Boltzmann Solvent Area (MM/PBSA) method. Here we show that ensemble simulations can correctly rank the binding affinities for these systems: we report the successful ranking of a drug binding to a variety of EGFR sequences, and of the two drugs binding to a given sequence.**

I. Introduction

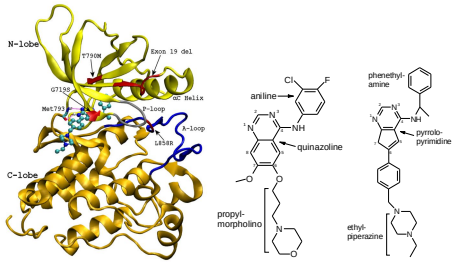
DRUGs used in "targeted therapy" are designed to identify and attack a specific molecular target, usually a protein which plays a critical role in tumor growth. As only specific targets are interfered with, targeted therapies are more effective and less harmful than most current combination treatments.

The EGFR is an especially important enzyme target in lung cancer because it mutates and/or is overexpressed in most non-small cell lung carcinoma (NSCLC) tumors. Tyrosine kinase inhibitors (TKIs) are frequently used to target the TK domain (Fig. 1a) and suppress its functions. Clinical studies manifest a strong correlation between the presence of mutations and patient response to tyrosine kinase inhibitors. A better understanding of the reasons for the success or failure of a therapeutic intervention will help in the selection of subgroups of patients who are most likely to respond to specific drugs, and pave the way for individualized treatment.

Manuscript received May 5, 2010. This research was funded by the EU FP7 ContraCancrum project (ICT-2007.5.30). Our work was partially supported by the National Science Foundation under NRAC grant MCA04N014, which provided access to Ranger at the Texas Advanced Computing Center (TACC), and by the DEISA Consortium (co-funded by the EU, FP7 project 222919) for access to HECToR and to Huygens. We also wish to acknowledge the UK NGS for providing access to their resources and their support for this project.

S. Wan is with the Centre for Computational Science, Chemistry Department, University College London, London, UK (e-mail: shunzhou.wan@ucl.ac.uk).

P. V. Coveney is with the Centre for Computational Science, Chemistry Department, University College London, London, UK (corresponding author to provide phone: +44(0)2076794850; fax: +44(0)2076791501; e-mail: p.v.coveney@ucl.ac.uk).



a) b) c)

Fig. 1. a) The intracellular tyrosine kinase domain of the epidermal growth factor receptor (EGFR) in complex with the ATP-competitive inhibitor AEE788 (Protein Data Bank (PDB) id: 2J6M). EGFR is depicted in ribbon and AEE788 in ball-and-stick representation. The principal elements used in this study are labelled. The kinase mutations are indicated in red. The chemical structures of Gefitinib and AEE788 are shown in b) and c) respectively, with the atom/group names and numbers used in this study.

II. Methods

Large scale molecular dynamics (MD) techniques are invoked to study the interactions between TKIs and EGFRs in atomistic detail, specifically employing MM/PBSA methodology [2] to predict the effect of mutations on drug binding affinities. The inhibitor binding free energy can be evaluated, following the MM/PBSA approach, as:

$$\begin{aligned}\Delta G_{binding} &= \langle \Delta E_{calc} \rangle - T\Delta S = \langle \Delta E_{MM} \rangle + \langle \Delta G_{solv} \rangle - T\Delta S \\ &= \langle \Delta E_{vdW} \rangle + \langle \Delta E_{elec} \rangle + \langle \Delta E_{PB} \rangle + \langle \Delta E_{SA} \rangle - T\Delta S\end{aligned} \quad (1)$$

Here, $\langle \Delta E_{calc} \rangle$ denotes the average of the calculated energy difference upon binding, which can be decomposed into molecular mechanics energy $\langle \Delta E_{MM} \rangle$ and solvation free energy $\langle \Delta G_{solv} \rangle$, and further into van der Waals $\langle \Delta E_{vdW} \rangle$, electrostatic $\langle \Delta E_{elec} \rangle$, electrostatic solvation free energy $\langle \Delta E_{PB} \rangle$ and nonpolar solvation free energy $\langle \Delta E_{SA} \rangle$. $-\langle T\Delta S \rangle$ term denotes the average contribution of configurational entropy change $\langle \Delta S \rangle$ at temperature T.

Ensemble simulations [5] are performed to rank the binding affinities of inhibitors AEE788 and Gefitinib (Fig. 1b,c) to wild-type and four mutant EGFRs. Our studies are

facilitated by a highly automated workflow tool, the Binding Affinity Calculator (BAC) [6], to launch molecular simulations and free energy calculations on numerous high-performance computing (HPC) resources.

III. Results

A. Ranking of binding free energies

In Fig. 2a, the calculated binding energies ΔE_{calc} (Eq. 1) are compared with the experimental data [1;4] for Gefitinib. Our calculated results are strongly correlated with the experimental values [1] in all cases except the L858R mutation, as shown in Fig. 2a. The L858R mutation appears to be an outlier, since in another

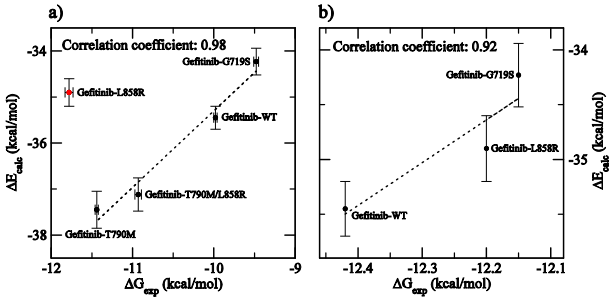


Fig. 2. Calculated binding energies (ΔE_{calc}) *vs* experimental binding free energies (ΔG_{exp}) for inhibitor Gefitinib with EGFRs. The calculated energies and their standard errors of the mean are obtained over all replicas in each ensemble simulation. The experimental data and their error bars are calculated from k_D values in reference[1] in (a) and from reference[4] in (b) (no error bars reported). The Gefitinib-L858R outlier is excluded in the linear fitting shown in (a).

experimental date set [4] it is well behaved (see Fig. 2b).

When combining calculated binding energies of AEE788 and Gefitinib and comparing with experimental data [1;3], a reasonable correlation is obtained by excluding three data points (Fig. 3a). A cross-drug correlation makes it possible to identify subgroups of patients who have a specific EGFR variant and are most likely to respond well to a particular drug treatment, and to choose a personalized drug therapy that maximizes treatment efficacy for an individual.

B. Determinant of free energy differences

The energy components (Eq. 1) analyses show that the electrostatic contribution (ΔE_{elec} + ΔG_{PB}) correlates well with the experimental binding free energies (Fig. 3b) while the non-polar contribution (ΔE_{vdW} + ΔE_{SA}) does not show any correlation, although the later is the most favourable component of the binding free energy. It is the electrostatic contribution which is largely responsible for the binding preferences of inhibitors to wild-type or mutant EGFRs.

C. Effect of mutations on drug binding affinities

To investigate what effects a mutation introduces to inhibitor binding, pairwise decomposition of the interaction energy is examined between the inhibitors and every residue of the receptor. The per-residue decomposition of the binding energy reveals how interactions change as a result of the mutations, and accounts for the molecular basis of drug efficacy.

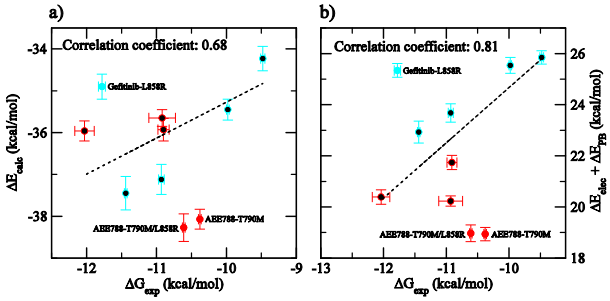


Fig. 3. Comparison of calculated binding energies (ΔE_{calc}) (a) and their electrostatic components (ΔE_{elec}+ΔE_{PB}) (b) with experimental binding free energies [1;3] (ΔG_{exp}) for inhibitors AEE788 (red) and Gefitinib (blue) complexed with EGFRs. Seven points (black) are used for linear fitting. The error bars are shown as standard errors of the mean from ensemble simulations and standard deviations from experiments [1;3]. Gefitinib-L858R is excluded for linear fitting as in Fig. 2. AEE788-T790M and AEE788-T790M/L858R are also excluded as they may have lower binding free energies as indicated in reference [7].

D. Reproducibility of free energy calculations

To validate the MM/PBSA method on the inhibitor-EGFR systems, ensemble simulations were performed with different initial structures. The energetic results are very similar, and not sensitive to the choice of initial structures, thanks to the ensemble simulation approach which renders conformational sampling considerably more extensive [8].

IV. Conclusions

The promising progress made in drug development for targeted cancer therapy invites the prospect of incorporating drug ranking into cancer decision support in order to predict drug sensitivity and resistance at the genotypic level. The ensemble method has the potential to accurately rank drug binding affinities on clinically relevant timescales (2-3 days), opening the way to their use in clinical decision support tools that match treatment to individual patients' genetic profiles [9].

References

[1] C. H. Yun, T. J. Boggon, Y. Q. Li, M. S. Woo, H. Greulich, M. Meyerson, et al., "Structures of lung cancer-derived EGFR mutants and inhibitor complexes: Mechanism of activation and insights into differential inhibitor sensitivity," *Cancer Cell*, vol. 11, no. 3, pp. 217-227, Mar.2007.

[2] S. Z. Wan, P. V. Coveney, and D. R. Flower, "Peptide recognition by the T cell receptor: comparison of binding free energies from thermodynamic integration, Poisson-Boltzmann and linear interaction energy approximations," *Phil. Trans. R. Soc. A-Math. Phys. Eng. Sci.*, vol. 363, no. 1833, pp. 2037-2053, Aug.2005.

[3] C. H. Yun, K. E. Mengwasser, A. V. Toms, M. S. Woo, H. Greulich, K. K. Wong, et al., "The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 6, pp. 2070-2075, Feb.2008.

[4] M. A. Fabian, W. H. Biggs, D. K. Treiber, C. E. Atteridge, M. D. Azimioara, M. G. Benedetti, et al., "A small molecule-kinase interaction map for clinical kinase inhibitors," *Nat. Biotechnol.*, vol. 23, no. 3, pp. 329-336, Mar.2005.

[5] S. K. Sadiq, D. W. Wright, O. A. Kenway, and P. V. Coveney, "Accurate Ensemble Molecular Dynamics Binding Free Energy Ranking of Multi-Drug-Resistant HIV-1 Proteases," *J. Chem. Inf. Model.*, in press. 2010. DOI:10.1021/ci100007w

[6] S. K. Sadiq, D. Wright, S. J. Watson, S. J. Zasada, I. Stoica, and P. V. Coveney, "Automated molecular simulation based binding affinity calculator for ligand-bound HIV-1 proteases," *J. Chem. Inf. Model.*, vol. 48, no. 9, pp. 1909-1919, Sept.2008.

[7] R. K. Kancha, N. von Bubnoff, C. Peschel, and J. Duyster, "Functional Analysis of Epidermal Growth Factor Receptor (EGFR) Mutations and Potential Implications for EGFR Targeted Therapy," *Clin. Cancer Res.*, vol. 15, no. 2, pp. 460-467, Jan.2009.

[8] L. S. D. Caves, J. D. Evanseck, and M. Karplus, "Locally accessible conformations of proteins: Multiple molecular dynamics simulations of crambin," *Protein Sci.*, vol. 7, no. 3, pp. 649-666, Mar.1998.

[9] S. K. Sadiq, M. D. Mazzeo, S. J. Zasada, S. Manos, I. Stoica, C. V. Gale, et al., "Patient-specific simulation as a basis for clinical decision-making," *Phil. Trans. R. Soc. A-Math. Phys. Eng. Sci.*, vol. 366, no. 1878, pp. 3199-3219, Sept.2008.

Web-Service Based Analysis of Gene-expression Data for Cancer Patients

J. Karlsson, M. Garcia, V. Martín-Requena, O. Trelles

***Abstract*— This paper demonstrates how web services provided in a secure environment can be used to efficiently process sensitive gene expression cancer data using a pipeline of web-services.**

I. Introduction

Data analysis in bioinformatics is typically delegated to external web-services which are designed to support interoperable machine-to-machine interaction over a network. The ACGT project [9] aims to develop methods and software systems which can improve the understanding of cancer research data through integration and analysis of biomedical information. Data analysis is performed in a controlled environment where user authenticity and data privacy is ensured.

The goal of this paper is to demonstrate how web services registered in the ACGT service repository [1] can be used to process gene expression data using a pipeline of several web-services.

II. Methods

A. Data set

The goal of the ACGT scenario "Multi-center multi-platform (MCMP)" is to show how data from two different microarray platforms (Affymetrix and Illumina) can be combined and analyzed. Biopsy samples of patients with breast cancer were collected in different medical center and the microarray technology was used to measure gene expression in the tumors [4]. This dataset contains demographic data and breast-cancer gene-expression in the format of single-color microarray data. Illumina data are synthetic, but based on real clinical data obtained with the Affymetrix technology.

B. Analysis methods

Two main platforms are common in gene expression analysis: *two-channel* (competitive) and *one-channel* hybridization. Two-channel microarrays are typically hybridized with cDNA prepared from two samples to be compared at the same time (e.g. diseased versus healthy samples), while one-channel (typically Affymetrix) contains a set of probes for genes (technical replications) from which probe-set values are obtained to represent the gene expression. Several analysis methods can be applied to ensure data quality, spot filtration, inter- and intra-slide normalization, replicate resolution, dye-swapping, random error removal and statistical analyses (see Prep [5]).

Once data is normalized, it can be used to generate a gene expression matrix. At this stage, more complex data analysis tools can be applied. One example is Engene [2], which is useful tool for storing, visualizing and processing large sets of expression patterns.

In the case of one-channel data, there are several alternatives to normalization of CEL data (for example, the Bioconductor R packages Affy [3]). These normalization methods can be applied in the ACGT platform with GridR services [10]. However, we are using a novel implementation of quantile normalization [7], which is well suited for large gene expression datasets.

Functionality from Prep, Engene and the quantile normalization implementation has been provided as ACGT compliant web-services [1] and registered in our metadata repository.

C. Analysis software

jORCA [6] (http://www.bitlab-es.com/jorca) is a powerful and portable desktop client with access to several service repositories with different protocols (including ACGT services), service discovery based on user data or keywords; embedded file system for handling local user files and user-defined lists of favorite tools.

jOrca is now able to enact pipelines constructed with automatic Web service composition (the user selects initial input data type and desired output data type and the system generates possible workflows). Workflows can also be created by selecting previous service invocations for

Manuscript received May 5, 2010. This work has been partially financed by the National Institute for Bioinformatics (www.inab.org) (Genoma España and the EC project "Advancing Clinico-Genomic Trials on Cancer" (Contract No.026996) and Fundación IMABIS/HCH Malaga.

J. Karlsson is with the Fundación IMABIS, Málaga, Spain (phone: +34 952137236; fax: +34 952132790; e-mail: tjkarlsson@uma.es).

M. Garcia (e-mail: maxgarcia@uma.es), V. Martín-Requena (e-mail: vickymr@uma.es) and O. Trelles (e-mail: ortrelles@uma.es) are with the Computer Architecture Department; University of Malaga, Málaga, Spain.

combination in a workflow.

III. Results

A. GE- data analysis with web-services

We have developed a large set of web-services for analysis of one-channel microarray data. Since normalization and processing of gene-expression data is strongly dependent on the quality of raw data and user evaluation, no unique "protocol" exists, but a set of individual procedures that must be combined to, collectively, produce the desired outcome. In this sense, in Figure 2, we show possible ways to combine the web-services.

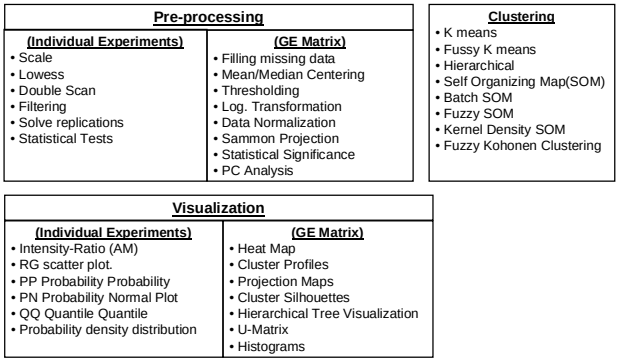


Fig. 1. This figure shows the different types of services currently available. Pre-processing services can be used to ensure data quality, spot filtration, normalization etc. Several clustering services are available when the data has been normalized, providing more advanced data analysis. The data can be visualized during the various steps. These services are possible to compose as workflows.

We describe some of those services below but for a full list of services, please see Figure 1 and also the page http://chirimoyo.ac.uma.es/gews.

- *QNormalization* [7]: This method assumes that there is an underlying common distribution of intensities across chips, using this assumption, different datasets can be given the same distribution by transforming the quantiles of each to have the same value by projecting on to the unit diagonal. Our implementation provides a fast procedure to normalize multiple large datasets gene-expression in parallel.
- *ReadOneChannel*: Produces a synthesized two-color micro array data by using raw gene expression microarray data (one channel) as input by creating a synthetic channel to obtain the pair control/target (using all the input files and applying average/median methods or using one or a set of microarrays files to calculate the channel).
- *HeatMAP* – Produces an image where gene expression values for each gene and experiment are visualized in a colored cell. The largest gene expression values are displayed in red (hot), the smallest values in green (cool) and unknown values in grey. The image size is fixed, and if the amount of data is high, some may not be represented.
- *Intensity-Intensity plot*: Produces an image where intensities of a set of slides are represented. This is a useful technique to observe the intensity deviation between slides before and after the normalization.

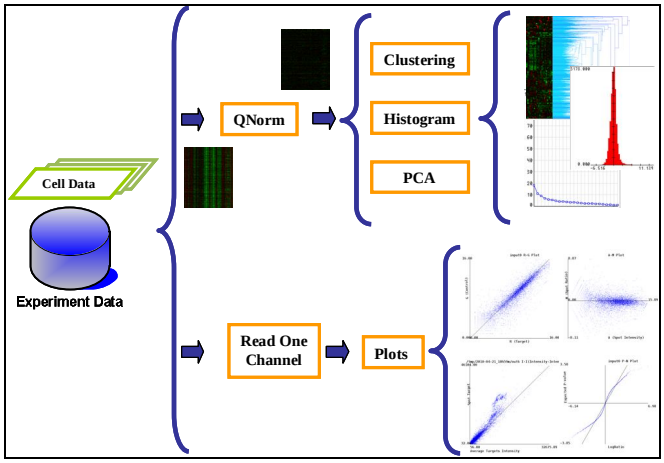


Fig. 2. The diagram shows different branches (possible workflows). For instance, in the top branch, data is normalized data using quantile method after which other methods (clustering, PCA, etc) can be performed. In the lower branch, one-channel cell data intensities is used to visualize all data experiments (ideally around bisect).

IV. Discussion and Conclusions

The pipelines presented in Figure 2 are basic but the goal of the paper is, rather than analyze the data *per se*, to demonstrate how web-services (developed within the ACGT environment) can be combined for data analysis of gene-expression data for tumors from cancer patients.

In Figure 3, we show how the data from Section II.A was pre-processed. We start with Affymetrix CEL-data, extract intensities with background correction without any normalization, and then visualize before and after QNormalization (inter-slide). This is, of course, a very initial attempt and by no means a complete analysis. It was, however, done as a proof-of-concept in a workflow created using jORCA. The possible steps for performing more complex analysis are shown in Figure 1. As can be observed, a large number of workflows are possible by using those services.

Of course, data protection is a priority when dealing with sensitive data. The approach in ACGT is to pseudonymize sensitive data (information which can be used for personal identification has been replaced with a label and association between labels and personal

information is only possible if the researcher has access to a mapping file). ACGT sets up an authority for data protection which enters into legally binding agreements with project participants to ensure data protection. This agency establishes a trusted third party which holds software tools and cryptographic keys, which can (when authorized) be used to de-anonymize data. All components in the architecture of ACGT require valid user credentials. Any data transfer between components, including the CLP services, is encrypted. This architecture is therefore very suitable for processing sensitive data.

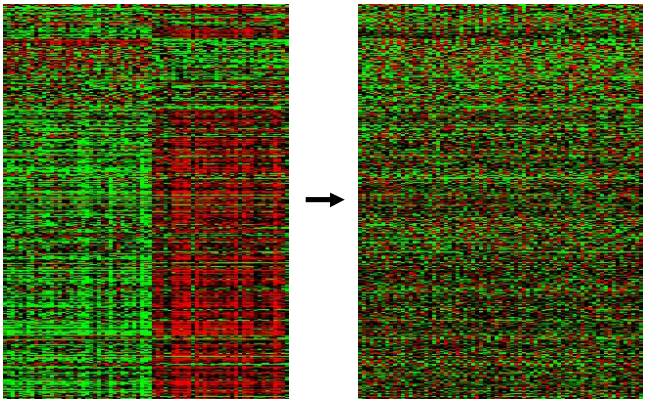


Fig. 3. To the left, the original MCMP data is visualized as a HeatMAP. To the right, the result of applying QNormalization to the data, visualized again as a HeatMAP.

There are several advantages with using external and distributed web-services. Since web-services are designed for re-use, their input and output data formats are publically described, thereby facilitating their combination in workflows. An additional advantage is that analysis with workflows can be reproduced by other groups. Input parameters to the workflows can also be varied to fine-tune the data analysis.

However, one challenge in using workflow is that they, by their very nature, depend on external resources (i.e. web services). A workflow can work perfectly fine when it is first developed but fail to run later because one (or several) of the web services in the workflow do not work.

This has not yet been addressed in this work but a possible solution could be automatic and regular testing of the functionality of web services and workflows to ensure continued functionality. However, such an approach does not guarantee that the service or workflow actually works in the moment when the user tries to use it. In [8], the possibility of service mirroring is proposed to allow several service instances to be installed and made available from different hosts in separate networks, providing redundancy.

Web services are useful but they have disadvantages, as explained earlier. However, a complex multi-task processing procedure can be made automatic using those web services to create useful and reproducible workflows. Privacy of patient data is ensured due to the secure ACGT architecture. In this work, we combine and integrate several techniques (web-services and workflows) through a user-friendly application (jORCA) and show a scenario of how those techniques can facilitate research processes related to analysis of gene expression data.

REFERENCES

[1] M. Garcia, et al: *Web service catalogue for Biomedical GRID infrastructure*, HealthGrid Conference 2010, accepted.

[2] J. Garcia de la Nava, et al: *Engene: the processing and exploratory analysis of gene expression data*. Journal of Bioinformatics 19(5), 657-658, 2003.

[3] R. Irizarry, et al. *An R package for analyses of affymetrix oligonucleotide arrays*. In The Analysis of Gene Expression Data: Methods and Software. Springer, New York, 2002.

[4] S. Loi, et al*: Definition of Clinically Distinct Molecular Subtypes in Estrogen Receptor–Positive Breast Carcinomas Through Genomic Grade*. Journal of Clinical Oncology 25(10), pp 1239-1246, 2007.

[5] V. Martin-Requena, et al: *PreP+07: improvements of a user friendly tool to preprocess and analyse microarray data*; BMC Bioinformatics 10(16), 2009.

[6] V. Martin-Requena, et al: *jORCA: easily integrating bioinformatics Web Services*, Journal of Bioinformatics 26(4), pp 553-559. 2010.

[7] J-M. Mateos-Duran, et al: *Qnorm: A library of parallel methods for gene-expression Q-normalization*. Proc. of BOSC 2009, p. 148. 2009.

[8] Ramirez, S., et al: *Mirroring Biomoby Services*. In: VII Spanish Bioinformatics Symposium. 2006.

[9] M. Tsiknakis, et al: *Building a European biomedical grid on cancer: the ACGT Integrated Project*. Studies in health technology and informatics, 120:247, 2006.

[10] D. Wegener, et al: *GridR: An R-based tool for scientific data analysis in grid environments.* In: Future Generation Computer Systems: The International Journal of Grid Computing: Theory, Methods and Applications, 25(4), 481-488, 2009.

Brain Energy Metabolism and Implications for Brain Tumors: An *In Silico* Study

Seda Genç[1], Işıl A.Kurnaz[2], Mustafa Özilgen[1]

***Abstract*—Glucose is known as the main and almost the only carbon source for cerebral energy metabolism. In recent years, there are some evidences that astrocytes perform glycolysis and lactate produced by astrocytes is taken up by neurons for aerobic respiration. In parallel, between astrocytes and neurons, a "lactate shuttle" hypothesis has been suggested. In the present study, two different models were constructed according to the classical view and the astrocyte-neuron lactate shuttle hypothesis (ANLSH) which suggests the presence of a lactate shuttle between astrocytes and neurons. These models were run under three conditions (normal glucose and normoxia; normal glucose and hypoxia; low glucose and normoxia), by using the biochemical simulation software package: COPASI 4.4.29. As a result, adenosine triphosphate (ATP) production in the second model (ANLSH) is greater than in the first model (classical view). The results of this study support the presence of the lactate shuttle between the astrocytes and the neurons.**

I. INTRODUCTION

Neuronal cells and astrocytes were thought to utilize glucose as an energy source for oxidative metabolism; however, the Astrocyte-Neuron lactate shuttle hypothesis (ANLSH) proposed in these past years predicts that glutamate recycling drives cerebral glucose metabolism and that glial glucose metabolism is almost completely anaerobic [1,2].

Lactate is a metabolite used also in hypoxia and normoxia in addition to anoxia, and lactate shuttle can be found in a variety of tissues including muscle, where there is, from muscle to the blood, a net flow of lactate which is then recovered from the blood by the resting muscle cell and removed from the system by oxidation [3].

This work was supported in part by the TUBİTAK with project number 107T380.

[1]Chemical Engineering Department, Yeditepe University, Kayışdağı, 34755, İstanbul, Turkey (phone: +90(216)578-04-88; fax:+90(216)578-04-00; e-mail:seda.genc@yeditepe.edu.tr, mozilgen@yeditepe.edu.tr).

[2]Genetics and Bioengineering Department, Yeditepe University, Kayışdağı, 34755, İstanbul, Turkey (e-mail: iakurnaz@yeditepe.edu.tr).

In this study, we have modeled both views separately and assessed their ATP production potentials. In the first model, the classical view assumes that both neurons and astrocytes can take up glucose and use it in glycolysis and aerobic respiraton (Fig.1). The pyruvate can choose two routes – some of it will be transported into mitochondria, converted into Acetyl Coenzyme A and enter the citric acid cycle, whereas some will be converted into lactate by lactate dehydrogenase (LDH) enzyme and secreted into the extracellular matrix through a generic monocarboxylate transporter, MCT (Fig.1).

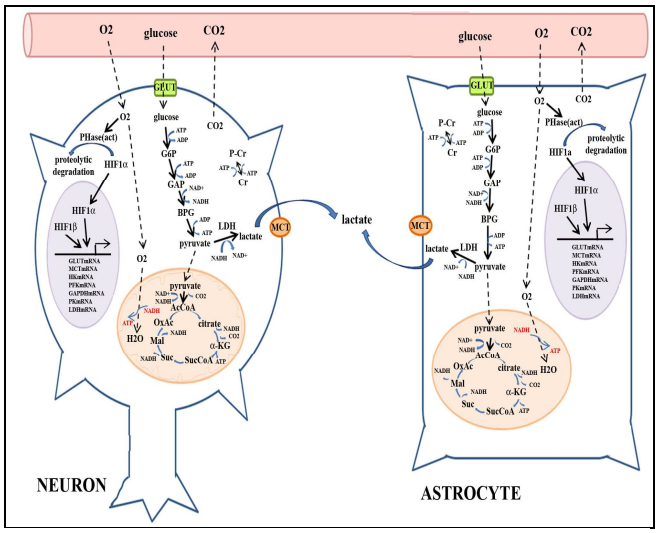


Fig.1. The classical view of energy metabolism within neurons and astrocytes

The second model, ANLSH, assumes that glucose is mainly taken up by the astrocyte and used up in glycolysis, the resulting pyruvate is converted into lactate by the astrocyte-specific LDH, and secreted out to the extracellular matrix via astrocyte-specific MCT. This lactate is in turn is taken up by the neuron via the neuron-specific MCT, and converted into pyruvate via neuron-specific LDH, which is then free to enter the citric acid cycle in mitochondria.

In both models, some of the key enzymes or transporters were modeled to be regulated in an oxygen-dependent manner through Hypoxia Inducible Factor [4,5]. Available oxygen levels are quite important for the survival of cells, and as such cells have devised methods to sense oxygen levels and respond accordingly.

This study focused on the energetic output of the classical view and ANLSH in the presence of hypoxia-dependent regulation of key enzymes. Our results show that the ANLSH is more advantageous for the neuron in terms of ATP produced, under both hypoxic and normoxic conditions, although it does not provide a significant advantage for the astrocyte.

II.MATERIALS AND METHODS

The biochemical simulation software package COPASI 4.4.29 was used for the analysis of our system. It is a freely-available program able to perform deterministic and stochastic simulation [6]. To simulate the metabolic processes that occur inside neuron and astrocyte during normoxia and hypoxia, a general mathematical model was developed where cells have interaction between capillary and extracellular area with distinct volume of nucleus, cytosol and mitochondrion domains. Between domains (nucleus-cytosol, cytosol-mitochondrion, and cytosol-extracellular area) molecular transport was assumed to occur either by passive diffusion (1) or carrier mediated transport (2).

	[Glu] (mM)	[O_2] (mM)		Classical View	ANLSH	Fold (shuttle/ classical)
NEURON	4.56	0.35	[ATP]_c	4.96	-	-
	4.56	7		4.85	-	-
	1	7		3.87	-	-
	4.56	0.35	[ATP]_m	128.81	497.43	3.86
	4.56	7		144	500.68	3.47
	1	7		131.35	382.18	2.90
ASTROCYTE	4.56	0.35	[ATP]_c	2.7	10.79	3.99
	4.56	7		2.7	10.79	3.99
	1	7		2.7	7.23	2.67
	4.56	0.35	[ATP]_m	149.11	-	-
	4.56	7		158.84	-	-
	1	7		145.20	-	-

Table 1. Comparison of cytoplasmic and mitochondrial ATP levels in neurons and astrocytes under different conditions

III.RESULTS

Our results indicate that under all three conditions studied (normal glucose and normoxia; normal glucose and hypoxia; low glucose and normoxia), ANLSH model provides on average the neuron with around 3-fold more mitochondrial ATP than under normoxia (Table 1).

Cytoplasmic ATP production in the astrocyte is also much more (around 2- to 4-folds) using the ANLSH; however it should be noted that in ANLSH it is assumed there is no mitochondrial ATP production, hence the overall astrocytic ATP production is significantly reduced. Oxygen and glucose deprivation (OGD) was previously shown to decrease the neuronal nicotinamide adenine dinucleotide (NADH) levels but not the astrocytic ones, and neurons were seen to be more susceptible to OGD-mediated cell death [7]. In the same study, it was shown that hypoxia was not detrimental to cells, but lack of glucose was more crucial – indeed in our simulations normoxia vs hypoxia does not change the levels of ATP significantly, whereas a decrease in glucose concentration has a serious negative effect.

IV.CONCLUSIONS

All tumors require high energy for their maintenance, and hypoxia or glucose deprivation in the tumorigenic area is a great problem, leading to increased blood vessel growth, or angiogenesis. Therefore, we believe that integrated computational models such as ours can be improved to represent the real-life situation in tumors cells, and in the future can be used to understand how to combat tumors of the brain, and even be used for drug-screening *in silico*.

REFERENCES

[1] L.Pellerin, and P.J.Magistretti, "Glutamate uptake into astrocytes stimulates aerobic glycolysis – a mechanism coupling neuronal-activity to glucose-utilization," *PNAS* 91, 10625–10629,1994.

[2] C-P.Chih, P.Lipton, Jr.E.L.Roberts, "Do active cerebral neurons really use lactate rather than glucose?," *Trends Neurosci*, 24**:**573–578,2001.

[3] L.B.Gladden, "Lactate metabolism: a new paradigm for the third millenium," *J* Physiol, 558:5-30, 2004.

[4] G.L.Semenza, "Hypoxia-Inducible Factor 1: Control of Oxygen Homeostasis in Health and Disease," Pediatr Res, 49:614–617, 2001.

[5] M.S.Ullah, A.J.Davies, A.P.Halestrap, "The plasma membrane lactate transporter MCT4, but not MCT1, is up-regulated by hypoxia through a HIF-1α-dependent mechanism", J Biol Chem, 281:9030–9037, 2006.

[6] S.Hoops, S.Sahle, R.Gauges, C.Lee, J.Pahle, N.Simus, M.Singhal, L.Xu, P.Mendes, U.Kummer, "COPASI - a COmlex PAthway Simulator," *Bioinformatics*, 22:3067–3074, 2006.

[7] A.Almeida, M.Delgado-Esteban, J.P.Bolanos, J.M.Medina, "Oxygen and glucose deprivation induces mitochondrial dysfunction and oxidative stress in neurons but not in astrocytes in primary culture," J Neurochem, 81:207-217, 2002.

Molecular Personalization of Cancer Treatment via a Multiscale Simulation Model of Tumor Response to Therapy. The Paradigm of Glioblastoma Treated with Temozolomide.

Amos A Folarin and Georgios S Stamatakos

***Abstract*—*In silico* modeling of tumor growth and treatment response within the ContraCancrum oncosimulator context requires a specific handling of the molecular variation of patient tumors which correspond to variable drug response. In this context we have looked to correlate the *in vitro* expression profiles with drug sensitivity in order to achieve a broad indication of sensitivity that can be used as a modifier of cell kill within the simulator. Using the exemplar case of the drug temozolomide in glioblastoma, we consider how a gene signature capable of predicting drug sensitivity can be utilized within a simulation of tumor growth and treatment response.**

I. Introduction

Glioblastoma multiforme tumors vary considerably in both histology and molecular biology. Several studies have shown the existence of multiple subtypes (mesencyhmal, proneural, classical and neural) classifiable by characteristic genes expressed in each subtype[1-2]. These subtypes have been shown to be of importance from a prognostic standpoint, with the highly aggressive and undifferentiated mesenchymal subtype of particular interest. However, the predictive capacity of these subtypes on drug response is less clear. Temozolomide, an alkylating agent that is able to cross the blood-brain barrier effectively is the current choice chemotherapeutic agent for treating glioblastoma. This drug is a member of a class of drugs that includes Dacarbazine (DTIC) and are prodrugs of MTIC (3-methyl-(triazen-1-yl)imidazole-4-carboxamide) which is able to alkylate the O-6 position of guanine residues. This chemical modification results through unknown mechanisms in DNA lesions during cellular DNA replication, followed by cell death triggered through the DNA damage response pathway. Response to temozolomide depends on tumor O6-methylguanine methyltransferase (MGMT) a protein which reacts stoichometrically to remove drug-induced alkylguanine adducts from DNA. Around half of patients with glioblastomas will not benefit from the drug because of effective MGMT-based repair of DNA base damage [3-4]. While MGMT is the dominant factor there are clearly additional pathways and genes are reportedly implicated in temozolomide resistance; these include the base excision repair (BER) pathway, in particular poly(ADP-ribose) polymerase 1(PARP1), and mismatch repair (MMR). Analogous to MGMT, the BER pathway plays an important role in repairing the cytotoxic methyl DNA adducts created by temozolomide, and high BER activity can confer tumor resistance to the effects of the drug. On the other hand, a deficiency in the MMR pathway can lead to DNA replication in spite of the methylated guanine residues thus diminishing the cytotoxic effects of temozolomide [5]. Further clues to temozolomide resistance are suggested by recent studies on stem or tumor initiating cells of glioblastoma with the presence of several *HOX* genes, inclusive of *Prominin-1* (CD133) segregating with poor survival in patients treated with concomitant temozolomide therapy[6].

Manuscript received May 17, 2010. This work was supported in part by the Clinically Oriented Translational Cancer Multilevel Modeling (CONTRA CANCRUM) project. Project number: FP7-ICT-2007-2- 223979.

A. A. Folarin is with the UCL Cancer Institute, Paul O'Gorman Building University College London, 72 Huntley Street, London, WC1E 6BT (corresponding author, phone: 303-555-5555; fax: 303-555-5555; e-mail: a.folarin@cancer.ucl.ac.uk).

G. S. Stamatakos is with the *In Silico* Oncology Goup, Institute of Communication and Computer systems, National Technical University of Athens, GR-157 80 Zografos, Greece (e-mail: gestam@central.ntua.gr).

II. Methods

At present, a mechanistic molecular model of patient specific drug resistance is not possible as many of the molecular components of both drug mechanism and tumor resistance are poorly defined. However, expression signatures which differentiate *in vitro* sensitivity to temozolomide have already been reported [7-8]. Our method is based on a comparison of a given patient's microarray expression profile to a gene expression model learned from *in vitro* drug response and microarray data. A classifier is trained from this using pertinent genes and classes defined by a drug sensitivity measure (e.g. GI_{50}).

For the first part we selected a database of several cancer lines screened against the drug temozolomide (we make the assumption that resistance mechanisms are conserved across these lines) and compare the genes that are differentially expressed between the most sensitive and resistant of these lines. From this set of genes, a support vector machine is trained and evaluated by its ability to accurately classify independent cell lines using cross-validation techniques.

III. Results

The known molecular profiles of the tumor type considered (e.g. glioblastoma multiforme) are classified from the treatment (e.g. temozolomide) responsiveness standpoint into either three groups (sensitive, intermediate and resistant) (Fig.1) or a more continuous set of sensitivity grades. This grouping is used in order to perturb the population based average values of the Cell Kill Ratio (CKR) or equivalently the Cell Kill Probability (CKP) or the Survival Fraction (SF) so that molecular personalization of the multiscale model is achieved. The quantitative extent of the perturbation is performed by starting with an empirically plausible fraction of the CKR (e.g. +1/3 CKR) which is to be added to CKR in the case of a sensitive tumor or to be subtracted from CKR (e.g. -1/3 CKR) in the case of a resistant tumor (Fig.2) and by subsequently applying an optimization loop . Obviously in the intermediate case no perturbation of the population based average value takes place. A similar strategy is applied to the more continuous molecular classification of tumor subtypes. The parameter of CKR (or equivalently CKP or SF) is of central importance to the cellular and higher biocomplexity level simulation since it provides an initial (although potentially very gross) estimate of the kill probability per individual (proliferating) cell of a clinical tumor [9-13].

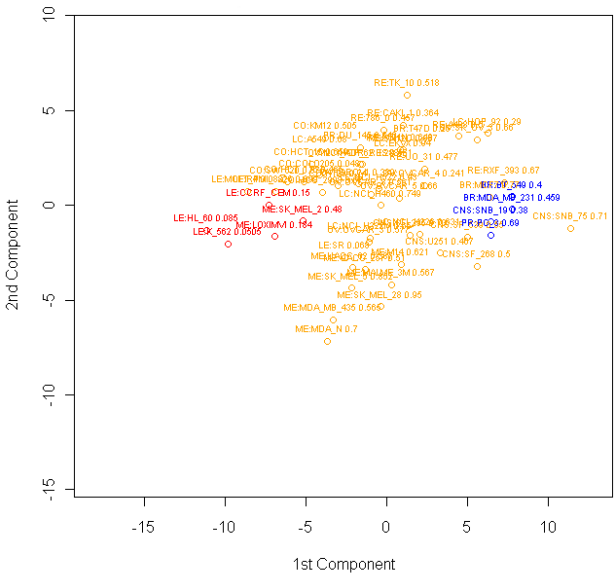


Fig.1. Segregation of the NCI60 database of cell lines along the first two principle components using the temozolomide gene signature. Classes determined from *in vitro* GI50 measures of sensitive (red), intermediate (yellow) and resistant (blue). GI50 stands for the concentration required to achieve 50 % growth inhibition.

IV. Conclusions

A brief outline of the exploitation of molecular data in order to molecularly personalize a multiscale model of clinical tumor response to chemotherapeutic treatment has been presented. The case of glioblastoma treated with temozolomide has been considered as a paradigm of the approach. A clinical adaptation and validation of the method described is in progress within the framework of the European Commission funded ContraCancrum project (: FP7-ICT-2007-2- 223979).

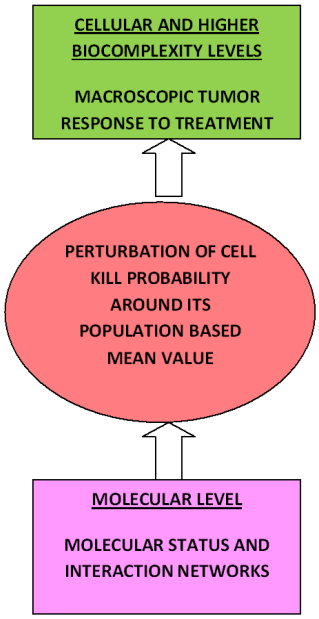


Fig.2. The *summarize and jump* strategy [12-13] as applied to patient individualization of treatment response

References

[1] Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD.et.al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell (2010) 17:98-110

[2] Phillips HS, Kharbanda S, Chen R, Forrest WF, Soriano RH, Wu TD.et.al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. Cancer Cell (2006) 9:157-173

[3] Hegi ME, Diserens AC, Gorlia T, Hamou MF, de TN, Weller M.et.al. MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med (2005) 352:997-1003

[4] Hegi ME, Liu L, Herman JG, Stupp R, Wick W, Weller M.et.al. Correlation of O6-methylguanine methyltransferase (MGMT) promoter methylation with clinical outcomes in glioblastoma and clinical strategies to modulate MGMT activity. J Clin Oncol (2008) 26:4189-4199
[5] Friedman HS, Kerby T, Calvert H Temozolomide and treatment of malignant glioma. Clin Cancer Res (2000) 6:2585-2597
[6] Murat A, Migliavacca E, Gorlia T, Lambiv WL, Shay T, Hamou MF.et.al. Stem cell-related "self-renewal" signature and high epidermal growth factor receptor expression associated with resistance to concomitant chemoradiotherapy in glioblastoma. J Clin Oncol (2008) 26:3015-3024
[7] Kitange GJ, Carlson BL, Schroeder MA, Decker PA, Morlan BW, Wu W.et.al. Expression of CD74 in high grade gliomas: a potential role in temozolomide resistance. J Neurooncol (2010)
[8] Yoshino A, Ogino A, Yachi K, Ohta T, Fukushima T, Watanabe T.et.al. Gene expression profiling predicts response to temozolomide in malignant gliomas. Int J Oncol (2010) 36:1367-1377
[9]G.S.Stamatakos, V.P.Antipas, and N.K. Uzunoglu , A spatiotemporal, patient individualized simulation model of solid tumor response to chemotherapy in vivo: the paradigm of glioblastoma multiforme treated by temozolomide , IEEE Transactions on Biomedical Engineering , 53(8) , 1467-1477 , 2006
[10] Stamatakos GS, Antipas VP, Uzunoglu NK. Simulating chemotherapeutic schemes in the individualized treatment context: the paradigm of glioblastoma multiforme treated by temozolomide in vivo. Comput Biol Med. 36(11) 1216-34 2006
[11] Dionysiou DD, Stamatakos GS, Gintides D, Uzunoglu N, Kyriaki K Critical parameters determining standard radiotherapy treatment outcome for glioblastoma multiforme: a computer simulation. Open Biomed Eng J (2008) 2:43-51
[12] G.S.Stamatakos, E.Kolokotroni, D.Dionysiou, E.Georgiadi, S.Giatili, *In silico* oncology: a top-down multiscale simulator of cancer dynamics. Studying the effect of symmetric stem cell division on the cellular constitution of a tumor. World Congress 2009 on Medical Physics and Biomedical Engineering, Sept 7-12, Munich, Germany. O. Doessel and W.C. Schlegel (Eds.): WC 2009, IFMBE Proceedings 25/IV, pp.1830-1833. www.springerlink.com 2009
[13] Stamatakos GS, Dionysiou DD Introduction of hypermatrix and operator notation into a discrete mathematics simulation model of malignant tumor response to therapeutic schemes in vivo. Some operator properties. Cancer Informatics (2009) 7:239-251

III. Cellular and Higher Level Based Modeling: Predominantly *Bottom-Up* Approaches

Application of ANOVA-Based Global Sensitivity Analysis to a Multiscale Cancer Model

Zhihui Wang and Thomas S. Deisboeck

***Abstract*—Using a two-dimensional multiscale non-small cell lung cancer model we present an analysis of variance (ANOVA) based global sensitivity analysis (GSA) method. With this novel method, we examine the cross-scale effects of molecular parameters on two tumor growth evaluation indices, tumor volume and expansion rate, at the multi-cellular level, by simultaneously varying parameter values over their entire parameter space. Results show that ERK, a downstream molecule of the EGFR signaling pathway, has the most important impact on regulating both the tumor volume and its expansion rate.**

I. Introduction

MULTISCALE cancer models, have begun to play a more important role in moving the field of integrative cancer systems biology towards clinical implementation [1]. However, to date, model parameters defining biological properties at different scales are not produced by a single laboratory, but instead are currently obtained from the literature or have to be estimated. It is thus crucial to not only study the dynamical system behavior governed by a *fixed* set of parameters, but to also further investigate the influence of their perturbations on the system outcome.

Sensitivity analysis has been widely accepted as a useful tool for identifying critical parameters to system output [2]. Global sensitivity analysis (GSA), a type of technique that addresses model behavior over a wide range of parameter operating conditions [3], can explore possible non-linear effects of the parameters on model outputs. Only recently have GSA methods been applied to systems biology models [4], most of which still focus on signaling pathway analysis. To successfully identify critical parameters in cancer models across different scales, a GSA method should be able to (1) deal with parameter uncertainties and simultaneous variation in inputs, (2) handle nonlinear input-output relationships, and (3) provide quantitative cross-scale measures of sensitivity.

In this article, we present a method based on analysis of variance (ANOVA) to perform GSA on a multiscale non-small cell lung cancer (NSCLC) model [5]. ANOVA is a model-independent method without assumptions regarding the functional relationships between input and output. This feature is of particular interest because, in most cases, such relationships are not known *a priori* or cannot be quantified mathematically. We believe that this is the first ANOVA-based systematic sensitivity analysis application to the theoretical modeling of cancer.

Manuscript received May 14, 2010. This work has been supported in part by NIH grant CA 113004 and by the Harvard-MIT (HST) Athinoula A. Martinos Center for Biomedical Imaging and the Department of Radiology at Massachusetts General Hospital.

Z. W. and T. S. D. are with the Harvard-MIT (HST) Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA 02129, USA (correspondence should be addressed to deisboec@helix.mgh.harvard.edu; phone: +1 617-724-1845; fax: +1 617-726-7422).

II. Methods

A. Multiscale Non-Small Cell Lung Cancer Model

We only briefly introduce key development methods of the 2D agent-based NSCLC model [5], which spans both molecular and multi-cellular scales. At the molecular scale, an epidermal growth factor (EGF)-induced, EGF receptor (EGFR)-mediated signaling pathway is implemented. At the multi-cellular scale, a lattice-based 2D microenvironment is constructed and populated with diffusive chemical cues including EGF, glucose, and oxygen. Figure 1 shows the schematic of the signaling pathway and the 2D tumor growth environment. As the simulation progresses, cancer cells constantly sense changes in environmental factors, interact with other cells, and adjust their behavior according to a set of predefined biological rules. Similar to previous studies [6-7], a cellular 'decision' algorithm is used to determine a cell's phenotypic transitions upon molecular changes: PLCγ-dependent migration and ERK-dependent proliferation.

B. ANOVA in Assessing a Parameter's Sensitivity

ANOVA evaluates the effects of the inputs on the response by decomposing the response into an overall mean (μ), main factor effects, interaction effects, and an error term (ε), and then provides their corresponding estimates [8]. In ANOVA, each input factor has to be

assigned specific ranges of values, i.e., factor levels. Here is an example of a two-way ANOVA with a single response variable (Y):

$$Y_{i,j,k} = \mu + A_i + B_j + (A \times B)_{i,j} + \varepsilon_{i,j,k} \quad (1)$$

where *i* refers to the level of factor *A*, *j* refers to the level of factor *B*, and *k* refers to the *k*th value of the response variable. A_i, B_j, and $(A \times B)_{i,j}$ represent the main effect of the *i*th level of factor *A*, the main effect of the *j*th level of factor *B*, and the interaction effect between the two factors. Corresponding equations can be deduced in the same way for three or more input factors. ANOVA uses an *F* value (obtained from *F*-tests) as the sensitivity index to evaluate the effects of a factor (input parameter) on model output

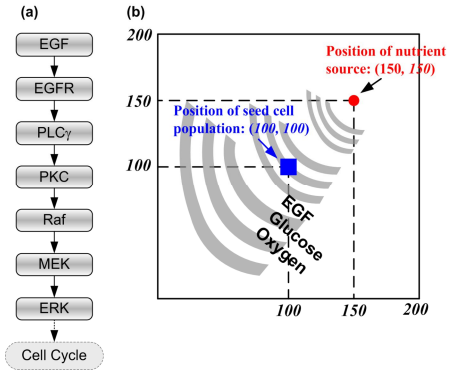


Fig. 1. (a) Schematic of the implemented EGFR signaling pathway within each cancer cell and (b) 2D biochemical microenvironment these cells face.

and then rank the parameters. The higher the *F* value, the more a parameter contributes to changes in the output, and thus the more *critical* the parameter is in the model.

III. Results

We only considered the initial concentrations of the pathway's seven molecules (Fig. 1) as input parameters. We created 2000 random sets of input parameter, i.e., 2000 simulation runs were carried out. The number of elapsed time steps was used as a measure for tumor expansion rate, and the final number of live cells for tumor volume.

Table I summarizes the ranking results. For tumor volume evaluation, ERK appears to be the most critical of the seven inputs, which is not surprising since ERK decides a cell's proliferation fate in our phenotypic decision algorithm of the 2D multiscale model [5]. MEK holds the 2nd position, but its *F* value is less than 10% of that of ERK, implying that it is far less important than ERK in influencing tumor volume. The remaining five molecules all have *F* values that are small. For the tumor expansion rate evaluation, ERK again is the top input, followed by PLCγ and EGFR. We assumed that PLCγ would have the most significant impact on tumor

TABLE I
ANOVA-BASED RANKING RESULTS

Rank	Input	*F* value
Output: Tumor volume		
1	ERK	236.52
2	MEK	15.5
3	Raf	9.02
4	PLCγ	7.72
5	PKC	6.4
6	EGF	4.91
7	EGFR	4.04
Output: Tumor expansion rate		
1	ERK	106.08
2	PLCγ	102.91
3	EGFR	90.4
4	MEK	15.43
5	EGF	7.19
6	Raf	6.65
7	PKC	6.39

expansion because it is the determinant of cell migration fate [5], but ERK turns out to be more important. The differences in *F* values for the top three parameters are not substantial, but the *F* value starts to drop dramatically from EGFR to MEK (the No. 4 parameter), indicating that the last four parameters (including MEK) are probably not as critical as the top three.

IV. Discussion

We analyze a multiscale NSCLC model to illustrate how ANOVA can be applied as a GSA method. ERK, the decision molecule for cell proliferation, appears to have the most significant impact on both of the selected tumor growth indices. It is counterintuitive that ERK is more important than PLCγ in regulating tumor expansion rate because PLCγ is used as the key component for determining a migratory phenotypic switch. However, it has been observed experimentally that the activation of ERK signaling is directly associated with both cancer growth and metastasis in many cancer types, including lung cancer [9]. Our analysis confirms thus theoretically the close relationship between ERK signaling and the final tumor outcomes.

Each sensitivity analysis method has its own assumptions, limitations, and demands regarding the time and effort needed for application and interpretation [10]. Hence, in identifying inputs critical, it is reasonable to consider the outcome of multiple sensitivity analysis methods. Therefore, in the future, to improve our understanding of such multiscale models, we plan on taking into account multiple GSA methods where a consistently high ranking of a parameter will indicate its overall importance.

References

[1] Z. Wang and T. S. Deisboeck, "Computational modeling of brain

tumors: Discrete, continuum or hybrid?," *Scientific Modeling and Simulation*, vol. 15, pp. 381-393, 2008.

[2] N. A. van Riel, "Dynamic modeling and analysis of biochemical networks: mechanism-based models and model-based experiments," *Brief Bioinform*, vol. 7, pp. 364-74, Dec 2006.

[3] S. Mishra, *et al.*, "Global sensitivity analysis techniques for probabilistic ground water modeling," *Ground Water*, vol. 47, pp. 730-47, Sep-Oct 2009.

[4] S. Marino, *et al.*, "A methodology for performing global uncertainty and sensitivity analysis in systems biology," *J Theor Biol*, vol. 254, pp. 178-96, Sep 7 2008.

[5] Z. Wang, *et al.*, "Simulating non-small cell lung cancer with a multiscale agent-based model," *Theor Biol Med Model*, vol. 4, p. 50, Dec 21 2007.

[6] Z. Wang, *et al.*, "Cross-scale sensitivity analysis of a non-small cell lung cancer model: linking molecular signaling properties to cellular behavior," *Biosystems*, vol. 92, pp. 249-58, Jun 2008.

[7] Z. Wang, *et al.*, "Cross-scale, cross-pathway evaluation using an agent-based non-small cell lung cancer model," *Bioinformatics*, vol. 25, pp. 2389-96, Sep 15 2009.

[8] B. J. Winter, *et al.*, *Statistical Principles in Experimental Design.*, 3rd ed. New York: McGraw-Hill Inc, 1991.

[9] J. Ming, *et al.*, "PRL-3 facilitates angiogenesis and metastasis by increasing ERK phosphorylation and up-regulating the levels and activities of Rho-A/C in lung cancer," *Pathology*, vol. 41, pp. 118-26, Feb 2009.

[10] H. C. Frey and S. R. Patil, "Identification and review of sensitivity analysis methods," *Risk Anal*, vol. 22, pp. 553-78, Jun 2002.

Tumor growth law and its physical and therapeutic implications

C. Guiot, P. Castorina, P.P. Delsanto, T.S. Deisboeck

***Abstract*—Therapeutic progress is based on a deeper understanding of the basic biological and physical-physiological features of tumor growth. The relevant contribution of the modeling approach to therapies is well documented by the dramatic improvement of chemotheraphy following the 'Norton-Simon' scheduling, based on the assumption of a 'Gompertzian' tumor growth curve, and by the proposal of many new therapies aiming at breaking off the energetic interplay between the tumor and its host by stopping angiogenesis and the local diffusion of various tumor-induced factors.**

Efforts to devise more realistic and biologically well-posed models will play a fundamental role towards the development of more effective 'individualized' therapies.

I. Introduction

Basic physical laws can be applied to tumor growth, and analogies with other physical systems may be suggestive for new approaches, aiming at a better understanding and treatment of cancer.

Tumor development, from its very beginning (exponential-like growth) to the late phases, exhibiting neoangiogenesis, invasiveness and metastasis, can be effectively described by means of a *'universal law'* (1) whose series expansion (2) generates successive terms accounting for the various phenomena listed above.

$$\dot{y}(t) = a(y,t)\,y(t), \qquad (1)$$

$$\dot{a} = \sum_{n=0}^{N} \alpha_n a^n = \beta a + \gamma a^2 + \ldots \qquad (2)$$

Manuscript received May 17, 2010. This work was supported in part by Regione Piemonte, PSF 2008 and University of Torino, Italy.

C. Guiot is with the University of Torino, Neuroscience Dept, C. Raffaello 30, 10125 Torino, Italy.; phone: +39116708166; fax: +39116708174; e-mail: caterina.guiot@unito.it.

P. Castorina , is with Università di Catania, Physic Dept, And Istituto Nazionale di Fisica Nucleare, Catania, Italy (e-mail: paolo.castorina@ct.infn.it).

P.P. Delsanto is with Politecnico di Torino, Physics Dept, Torino, Italy (e-mail: pier.delsanto@polito.it).

T.S. Deisboeck is with Harvard-MIT (HST) Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, USA
(e-mail: DEISBOEC@HELIX.MGH.HARVARD.EDU)

Such 'universal law' can be related to basic physical parameters and to energy conservation principles [1]. For instance, the first three terms correspond to exponential (U0), Gompertzian (U1) [2] and 'West-like' (U2) [3] growths respectively. 'West-like' growth, in particular, includes the fractal topological patterns observed both in tumor invasion branching processes (driven by the interplay between tumor surface tension and cohesive forces) and in neoangiogenesis (see Fig.1).

In biological terms, stages U1 and U2 account for the ability of tumors to modify their metabolic substrates, overcoming tissue hypoxia and to colonize the host organ by expressing a variety of factors promoting an effective interplay.

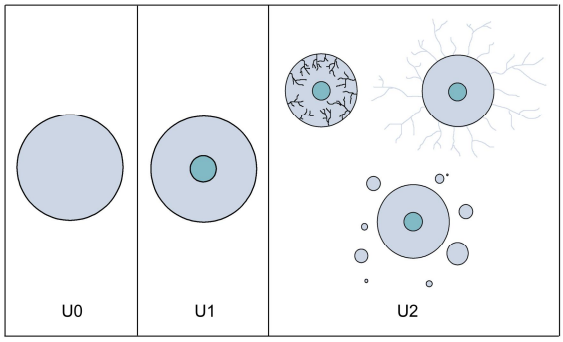


Fig. 1. Schematic representation of the three UN phases. In U0 the experimental tumor has no core, but it starts developing it in U1. The third phase (U2) represents three possible evolutionary scenarios of a tumor "in vivo", i.e. with angiogenesis, tumor invasion and metastatic dissemination.

As a matter of fact, conventionally, therapy is based on a 'naïve' description of tumor growth, i.e. its main target is stopping cell division of the (concomitantly) proliferating cells, which is very effective at stage U0 but far less useful at later stages. Since therapies are mainly administrated when tumors have already developed effective strategies of interplay with the host (organ/body), only an approach focused on 'isolating the enemy before fighting' can be effective. Such therapeutic effectiveness (e.g. radiotherapy and adjuvant chemotherapy schedules) would therefore greatly benefit from a more physical description of tumor growth, e.g. on the basis of the 'universal' growth law.

II. IMPLICATIONS FOR THERAPY

As widely described in the literature, the assumption of a Gompertzian (U1) growth law, instead of simpler U0 models, is the basis for the so-called ‘ Norton-Simon chemotherapy schedule’[4,5] , which dramatically improved chemotherapy effectiveness by showing that not only overall drug dose, but also its ‘density’, i.e. its administration at shorter intervals, can add benefit in combating cancer cell regrowth. To predict the improvements obtained by administrating ‘drug cocktails’ adjuvant therapies , a simple model [6] has been recently proposed, where tumor cells, after an initial undifferentiated phase, are assumed to express different features belonging to two different ‘clones’, which are possibly sensitive to different drugs, and using a ‘two population dynamics’ approach. The model shows that when a small tumor cell fraction emerges which exhibits a higher proliferation rate than the parent strain, a tumor instability occurs. If this instability precedes the onset of treatment, the slope of the linear increase of the drug concentration required by the standard “Norton-Simon late intensity schedule” changes and the drug dose strongly depends on the balance of the two tumor cell populations and on their growth rates. Also if the instability occurs after the initial treatment, the effective chemotherapeutic drug concentration for tumor control changes as well. In both cases modifications to the “Norton-Simon late intensity” schedule should be performed to enhance therapeutic results.

The impact of the choice of the U2 growth model on treatment strategies has been investigated by simulating the effects of various radiotherapeutic regimen [7].

By comparing the tumor survival fraction during RT regimen predicted by the U1 and U2 growth laws, which are based on scaling principles, we noted that while in the first case the surviving tumor cell fraction could be reduced 'ad libitum', simply by increasing the number of radio-therapeutic fractionated treatments, independently of the initial tumor mass, U2 established a lower limit for the survival fraction, weakly dependent on the clonogenic number, that could not be reduced any further regardless of the total number of treatments. This could explain the clinical evidence for better responses of large tumors to some non-conventional schedules, such as hypofractionation and CHART, in comparison with conventional ones, since the U2 model was more appropriate to account for tumor regrowth of highly proliferating tumors during RT.

Finally, a recently published paper [8] investigates the evolution of prostate cancer from diagnosis to death by effectively combining a detailed physical description of tumor growth at different stages, a realistic, clinically based evaluation of the values of the model parameters (based on tumor volume measurements, PSA and Gleason score throughout the natural tumor history) and a very effective statistical inference procedure on previous clinical reports. As a result, a powerful algorithm capable of predicting disease progression and survival probability is derived which, provided further validation on large patient populations, could be the ‘reference’ for a new modeling approach aiming at testing innovative treatment strategies and ‘customizing’ therapies for individual patients, allowing ‘real time’ adjustments and for upgrade of treatments and drug doses.

III. CONCLUSIONS

We present new insights on the actual growth of tumors, both based on more realistic and properly parametrically based algorithms (see the ‘U(N) law’ described by Eqns (1,2)) that account for the complex relationships occurring between the tumor and its host throughout carcinogenesis. On such basis the traditional, well consolidated therapeutic strategies can be critically revised and doses and schedules can be ‘optimized’ and ‘customized’ according to different lesions and patient needs.

On the other hand, a better understanding of the physical and basically ‘energetic’ role of the complex molecular machinery relating the developing cancer with its surrounding may help devising new therapeutic strategies.

REFERENCES

[1] P. Castorina PP. ,Delsanto , C.,Guiot C .”Classification scheme for phenomenological universalities in growth problems in physics and other sciences.”. *Phys Rev Lett* , vol 96, May 2006, pp. 188701-.

[2] 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.,* vol 123, 1825, pp. :513- .

[3] GB. West, JH Brown , BJ Enquist.”A general model for ontogenetic growth” . *Nature*, vol. 413, Oct 2001, pp. :628-631.

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

[5] L. Norton. “Conceptual and practical implications of breast cancer geometry: Toward a more effective, less toxic therapy”. *The Oncologist* vol. 10 , Jun-Jul 2005, pp. 370-381.

[6] P. Castorina, D. Carcò, C. Guiot, TS. Deisboeck. “ Tumor growth instability and its implications for chemotherapy”. *Cancer Res*. November 2009, pp. 8507-8515.

[7] P. Castorina P., TS.Deisboeck ,P. Gabriele P., C.Guiot. “Growth laws in cancer: implications for radiotherapy” *Radiation Res* vol 168, September 2007, pp. :349-356

[8] G. Dimonte. “A cell kinetics model for prostate cancer and its application to clinical data and individual patients “*J. Theoret. Biol* vol 264, May 2010, pp. 420-442Oct November 2009, pp. 8507-8515 November 2009, pp. 8507-8515.

Tumor microenvironment in a real-life model of tumor spheroids

Roberto Chignola and Edoardo Milotti

***Abstract*—Tumors are complex bio-systems and cell growth is coupled to the chemical modifications of the extracellular microenvironment. Tumor cells and their microenvironment, therefore, constitute an evolving cellular ecosystem and a detailed understanding of the underlying dynamics might provide insights into tumor development and resistance to therapy. Here we present a real-life computer program for the simulation of multicell tumor spheroids. Simulation results compare quite well with experimental data and yields unique view of tumor microenvironment.**

I. Introduction

INDIVIDUAL tumors are complex biological systems and, in spite of great therapeutic advances, many tumors still escape treatment and lead to death. Microscopically, the malignant transformation of single cells is a multistep process that involves the modification of several molecular circuits that, in turn, modify the cells' behavior and the relationships between cells and the environment [1]. In addition, epigenetic and environmental factors, which include cell-cell interactions, also conspire with the bare genetic information to make tumor growth a highly variable process with very strong feedbacks. Tumor cells and their microenvironment, therefore, constitute a complex cellular ecosystem and a detailed understanding of the laws that govern its evolution might provide novel insights into tumor biology and resistance to therapy.

Multicell tumor spheroids (MTS) display several biological characteristics observed in the inter-vascular regions of real tumors *in vivo*. These include, heterogeneous expression of adhesion molecules, production of an intercellular matrix, heterogeneous distribution of nutrients and oxygen that lead to central region of cell death [2]. In particular, MTS show resistance phenomena to anti-tumor therapies [3], and a marked individual response to chemotherapy [4] that parallels a great variability in individual MTS growth kinetics [5]. An explanation of these experimental observations is still missing.

The heterogeneous growth kinetics of individual MTS and their heterogeneous response to anti-tumor therapy are rather disturbing biological facts and reveal that, in spite of their apparent simplicity as an experimental model of tumors, MTS possess a hidden, yet unexplored, complexity where random fluctuations also play an all-important role. This complexity can only be explored with a fine-grained model of MTS.

Manuscript received April 20, 2010. This work was supported in part by grants from the Istituto Nazionale di Fisica Nucleare (I.N.F.N., group V, VBL-RAD).

R. Chignola is with the Dipartimento di Biotecnologie, Università di Verona, Strada Le Grazie 15 – CV1, I-37134 Verona, Italia, and I.N.F.N. Sezione di Trieste, Via Valerio 2, I-34127 Trieste, Italia (phone: +39-045-8027953; fax: +39-045-8027929; e-mail: roberto.chignola@ univr.it).

E. Milotti is with the Dipartimento di Fisica, Università di Trieste and I.N.F.N. Sezione di Trieste, Via Valerio 2, I-34127 Trieste, Italia (e-mail: edoardo.milotti@ts.infn.it).

II. A Real-Life Model of MTS

We have previously modeled the basic features of living tumor cells and these include the main metabolic and biosynthetic pathways, the control of the cell cycle and cell death [6,7]. This model has been tested for its ability to simulate the metabolism and the growth of dispersed tumor cell populations, and simulation outputs have been compared to actual data on a quantitative basis [6,7]. Recently, we have extended the model to include other biochemical details and the extracellular environment.

The present MTS model is based on this working model of tumor cells. A cell takes up nutrients from the extracellular environment and releases into it its waste products, such as lactic acid. This in turn decreases the environmental pH and affects the capacity of molecular transporters to absorb nutrients. In this way, the model realizes a nonlinear feedback between cells and their surroundings.

The MTS model has a full 3-dimensional geometry and a cell interacts with the others by means of biomechanical forces. It is therefore essential to keep track of neighboring cells, and the proximity relations among cells are provided by a Delaunay triangulation [8]; the triangulation also serves as the scaffolding to solve the diffusion problem for each biochemical species within the MTS structure. It is important to note that the MTS model is off-lattice and cells can move freely in the space under the action of the forces given by the pushing of dividing cells and the

pulling of dying cells whose volume shrinks.

All model parameters are derived from experiments or deduced from reasonable theoretical arguments, so that the model does not contain free parameters. Parameters can only vary within biophysically meaningful ranges and the model is thus truly predictive. It provides qualitative behaviors but also quantitative outputs, so that simulations can be straightforwardly compared to actual data

III. Results

A number of experimental observations carried out on individual MTS are available for comparison purposes with simulation outputs. These include: growth kinetics, where MTS size is measured over time [4,5], distribution of nutrients, lactic acid and energy molecules such as ATP as a function of MTS radius [2,9], pH gradient within the MTS [2,10], size of the necrotic core [11] and many others.

Simulation outputs turn out to be in good quantitative agreement with experimental data as far as the growth kinetics, metabolic and morphologic parameters are concerned. For example, a simulated MTS of approximately 500 μm diameter showed a viable rim thickness of 155 μm and a hypoxic rim thickness of 98 μm (140-300 μm and 50 μm in actual experiments, [10,11]), a central pO2 of 7 mmHg (0-20 mmHg n real MTS, [9]) a central pH of 6.7 (6.6-7.0 in actual data,

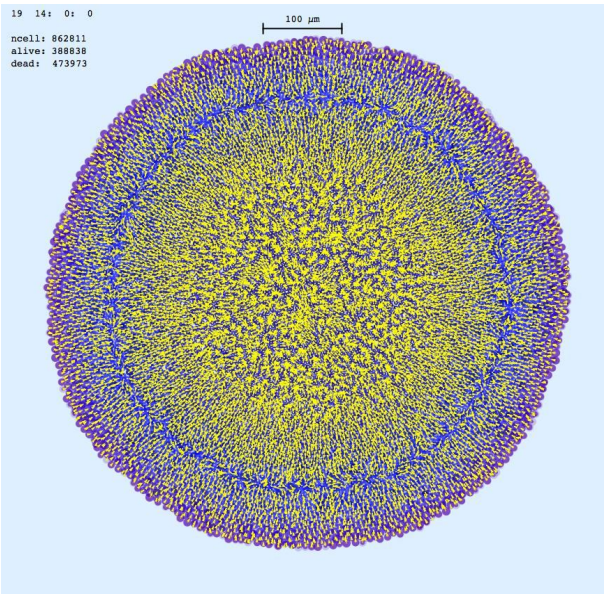


Fig. 1. Flow of glucose (arrows) along the central section of a MTS superposed to glucose concentration (given in pseudo colors: blue low, red high). Numbers in the top left corner show real time, so that this MTS is 19 days and 14 hours old.

[10,12]). Most importantly, however, it should be recalled that these (and other) values have not been estimated by fitting model equations to experimental data. Once parameters are fixed, a singe cell is allowed to evolve up to a population of more than 10^6 cells, and model outputs are therefore true predictions of MTS biology.

Since cellular niches are continuously updated as cells proliferate and die, the model provides unique dynamic view of tumor microenvironment. We observe chaotic motions of the cells and self-organizing flows of nutrients within the MTS structure. An example is given in Fig.1 that shows the flow of glucose in the MTS. In the outer layers the flow is inward whereas in the central core is outward. We expect a similar flow for water, though water is not presently considered in the model. If this is so, then the water outflow would hinder the inward diffusion of slowly moving molecules, and this might provide a biophysical explanation for MTS resistance to therapy.

IV. Conclusion

Our MTS model gives outputs in good quantitative agreement with experimental data. The model provides true *ab initio* predictions of MTS evolution at spatiotemporal resolutions well below the limits of present experimental settings. The model looks like quite robust to parameter changes, but its robustness will be demonstrated by a sensitivity analysis in the next future. Since the model is partly stochastic, the effects of fluctuations and the extent of correlations between model variables will also be investigated by means of an extensive simulation campaign that is already programmed for the next months. However, the individual runs carried out so far give new hints into the evolution of MTS and their resistance to chemotherapy.

References

[1] D. Hanahan, and R. A. Weinberg, "The hallmarks of cancer", *Cell*, vol. 100, pp. 57-70 (2000).

[2] R. M. Sutherland, "Cell and environment interactions in tumor microregions: the multicell spheroid model", *Science*, vol. 240, pp. 177-184 (1988).

[3] H, Kobayashi, S. Man, C. H. Graham, S. J. Kapitain, B. A. Teicher, and R. S. Kerbel, "Acquired multicellular-mediated resistance to alkylating agents in cancer", *Proc. Natl. Acad. Sci. USA*, vol. 90, pp. 3294-3298 (1993).

[4] R. Chignola, R. Foroni, A. Franceschi, M. Pasti, C. Candiani, C. Anselmi, G. Fracasso, G. Tridente, and M. Colombatti, "Heterogeneous response of individual multicellular tumor spheroids to immunotoxins and ricin toxin", *Br. J. Cancer*, vol. 72, pp. 607-614 (1995).

[5] R. Chignola, A. Schenetti, G. Andrighetto, E. Chiesa, R. Foroni, S. Sartoris, G. Tridente, and D. Liberati, "Forecasting the growth of multicell tumor spheroids: implications for the dynamic growth of solid tumors", *Cell Prolif.*, vol. 33, pp. 219-229 (2000).

[6] R. Chignola, and E. Milotti, "A phenomenological approach to the

simulation of metabolism and proliferation dynamics of large tumor cell populations", *Phys. Biol*., vol. 2, pp. 8-22 (2005).

[7] R. Chignola, A. Del Fabbro, C. Dalla Pellegrina, E. Milotti, "Ab initio phenomenological simulation of the growth of large tumor cell populations", *Phys. Biol*., vol. 4, pp. 114-133 (2007).

[8] See, e.g., M. de Berg, O. Cheong, M. van Kreveld, and M. Overmars, *Computational geometry: algorithms and applications*. Berlin: Springer-Verlag, 2008, 3rd ed.

[9] S. Walenta, J. Doetsch, W. Mueller-Klieser, and L. A. Kunz-Schughart, "Metabolic imaging in multicellular spheroids of oncogene-transfected fibroblasts", *J. Hisochem. Cytochem*., vol. 48, pp. 509-522 (2000).

[10] H. Acker, J. Carlsson, G. Holtermann, T. Nederman, and T. Nylen, "Influence of glucose and buffer capacity in the culture medium on growth and pH in spheroids of human thyroid carcinoma and human glioma origin", *Cancer Res*., vol. 47, pp. 2504-3508 (1987).

[11] R. M. Sutherland, B. Sordat, J. Bamat, H. Gabbert, B. Bourrat, and W. Mueller-Klieser, "Oxygenation and differentiation in multicellular spheroids of human colon carcinoma", *Cancer Res*., vol. 46, pp. 5320-5329 (1986).

[12] J. Alvarez-Perez, P. Ballesteros, and B. Cardan, "Microscopic images of intraspheroidal pH by 1H magnetic resonance chemical shift imaging of pH sensitive indicators", *MAGMA*, vol. 18, pp. 293-301 (2005).

Physical and Computational Issues in a Simulation of Multicellular Tumor Spheroids

Edoardo Milotti and Roberto Chignola

***Abstract*—We have developed a computer program which simulates the growth and development of multicellular tumor spheroids. The program implements a basic description of the metabolism, growth and proliferation of single cells, a full 3-dimensional geometry, and handles the complex problem of diffusion and transport of nutrients and metabolites into and out of cells, and in their surrounding environment. Here we discuss some of the challenging computational problems that arise in the implementation of this biophysical model.**

I. Introduction

We have developed a complex simulation program that implements a numerical model of multicellular tumor spheroid (MTS) growth and development. The program structure is based on a model of single cells, documented in a series of papers [1-3], and is shown schematically in figure 1.

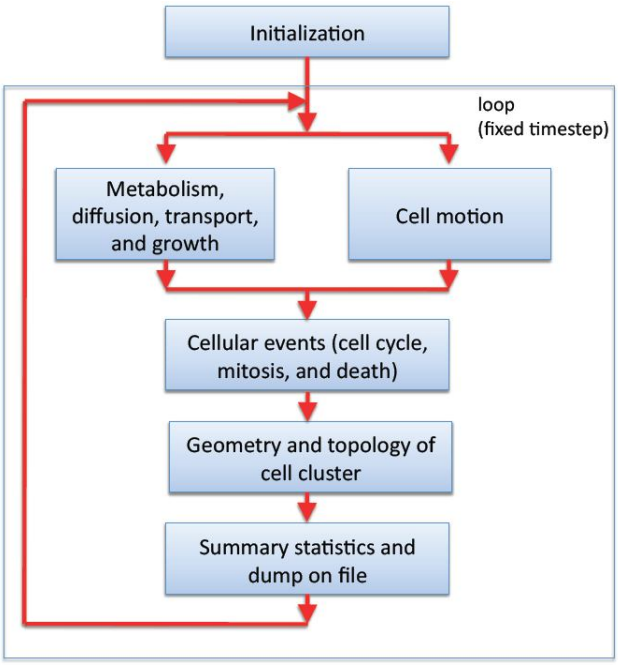


Fig. 1. Schematic diagram of the program structure.

After initialization, the program enters a loop that computes the state of the system of cells at fixed time steps. The loop calculates the evolution of the biochemical variables of each cell ("Metabolism, diffusion, transport, and growth" step) and the mechanical motions of the cell cluster ("Cell motion"). These two steps can proceed in parallel, and at present the actual parallelization is performed in a SIMD framework using OpenMP [4], and either with the GNU [5] or the Intel [6] C++ compiler. The next step ("Cellular events") checks each cell in the cluster and carries out all necessary state changes; in case of mitosis this step also adds new cells to the list of cells. This is followed by a step that computes the geometrical and topological properties of the cluster: in particular, this step calculates the proximity relations of each cell, by way of a Delaunay triangulation [7], and the positions of cells that lie on the boundary of the cluster, by means of an alpha-shape calculation [8,9]. A final program step outputs relevant cluster statistics on file ("Summary statistics and dump on file").

II. Mathematical Treatment of Metabolism, Transport and Diffusion

We discretize the diffusion problem in the cell cluster using the cells themselves as the basis for discretization. In this way we obtain a disordered, discrete lattice of sites, and a large set of nonlinear differential equations that combine diffusion on this lattice – where each site represents the extracellular volume around each cell – with transport between each cell and the surrounding extracellular volume. This means that each cell has two spatial compartments (the inside of the cell – denoted by uppercase subscripts in the equations – and by the extracellular space surrounding it – denoted by lowercase subscripts), and the generic form of the equations is

$$V_C \frac{d\rho_C}{dt} = M(\rho_C) + T(\rho_c, \rho_C), \tag{1a}$$

$$V_c \frac{d\rho_c}{dt} = -T(\rho_c, \rho_C) + D\sum_{\langle b \rangle}(\rho_b - \rho_c)g_{bc}, \tag{1b}$$

Manuscript received April 20, 2010. This work was supported in part by the Istituto Nazionale di Fisica Nucleare (I.N.F.N.).

E. Milotti is with the Dipartimento di Fisica, Università di Trieste and I.N.F.N. – Sezione di Trieste, Via Valerio 2, I-34127 Trieste, Italy (phone: +39-040-558-3388; fax: +39-040-558-3350; e-mail: edoardo.milotti@ts.infn.it).

R. Chignola, is with the Dipartimento di Biotecnologie, Università di Verona and with the I.N.F.N. – Sezione di Trieste, Strada Le Grazie 15 CV1, I-37134 Verona, Italy (e-mail: roberto.chignola@univr.it).

$$V_c \frac{d\rho_c}{dt} = -T(\rho_c, \rho_C) + D\sum_{\langle b \rangle}(\rho_b - \rho_c)g_{bc} + D_{env}(\rho_{env} - \rho_c)g_c ,$$

(1c)

$$V_{env} \frac{d\rho_{env}}{dt} = -D_{env}\sum_{\langle c \rangle}(\rho_{env} - \rho_c)g_c + (\rho_{in} - \rho_{env})f .$$

(1d)

where the ρ's denote concentrations. Equations like (1a) describe internal metabolism (the *M* term) and transport between cell and extracellular volume (the *T* term). Equations like (1b) describe the evolution of concentrations in the extracellular volume, which is associated to transport and simple diffusion (second term on the r.h.s.). Equations like (1c) hold for cells that are in contact with the external environment, so that there is an additional diffusion term that contains enviromental variables. Finally, equations like (1d) describe the evolution of environmental concentrations.

It can be shown that an algorithmically stable solution is obtained only with implicit integration methods [10], and that a simple method like implicit Euler integration yields satisfactory results [10].

Implicit integration of the equations leads in turn to the following set of nonlinear equations (where masses are used as well as concentrations)

$$m_C^{n+1} = m_C^n + \Delta t\left(M\left(m_C^{n+1}\right) + T\left(m_c^{n+1}, m_C^{n+1}\right)\right)$$

(2a)

$$m_c^{n+1} = m_c^n + \Delta t\left(-T\left(m_c^{n+1}, m_C^{n+1}\right) + D\sum_{\langle b \rangle}\left(\frac{m_b^{n+1}}{V_b} - \frac{m_c^{n+1}}{V_c}\right)g_{bc}\right)$$

(2b)

$$m_c^{n+1} = m_c^n + \Delta t\left(-T\left(m_c^{n+1}, m_C^{n+1}\right) + D\sum_{\langle b \rangle}\left(\frac{m_b^{n+1}}{V_b} - \frac{m_c^{n+1}}{V_c}\right)g_{bc} + D_{env}\left(\rho_{env} - \frac{m_c^{n+1}}{V_c}\right)g_c\right)$$

(2c)

$$\rho_{env}^{n+1} = \rho_{env}^n + \Delta t\left(-\frac{1}{V_{env}}\left(\sum_{\langle c \rangle} D_{env}\left(\rho_{env}^{n+1} - \frac{m_c^{n+1}}{V_c}\right)g_c\right) + \left(\rho_{in}^{n+1} - \rho_{env}^{n+1}\right)\frac{f^{n+1}}{V_{env}}\right)$$

(2d)

The nonlinear equations (2a-d) are solved with a combination of Newton-Raphson and secant method steps: this is further illustrated in figure 2.

The method discussed here turns out to be exceptionally robust, and we have applied a variant to the solution of the mechanical part as well. On the whole, at each timestep we integrate 25 coupled nonlinear differential equations per cell, and since our largest simulation to date actually exceeded 1.5 million cells, the integrator in the program has proved to be able to handle as many as 37.5 million coupled nonlinear differential equations.

III. Conclusion

We have developed a robust scheme for the integration of a very large number of coupled nonlinear differential equations such as those that regulate diffusion and transport in a large cluster of cells. The integration scheme is an essential part of our simulation of MTS, and has led to unprecedented views of their internal machinery [11]. Figure 3 shows an example image obtained in a preliminary run: the upper panel shows a central slice of a 16-day-old simulated spheroid, where the gray level maps the oxygen concentration. Oxygen concentration is high along the boundary with the environment and decreases fast towards in the central part of the simulated spheroid. We have full access to all metabolic variables in the programs and it is easy to produce a plot of oxygen partial pressure vs. the distance from the centroid of the simulated spheroid: this is shown in the lower panel of figure 3. A comparison of this and other similar plots with available experimental data [12-16] shows excellent agreement.

We conclude that the computational scheme that we have developed is reliable, robust, and has a large discovery potential that waits only to be tapped.

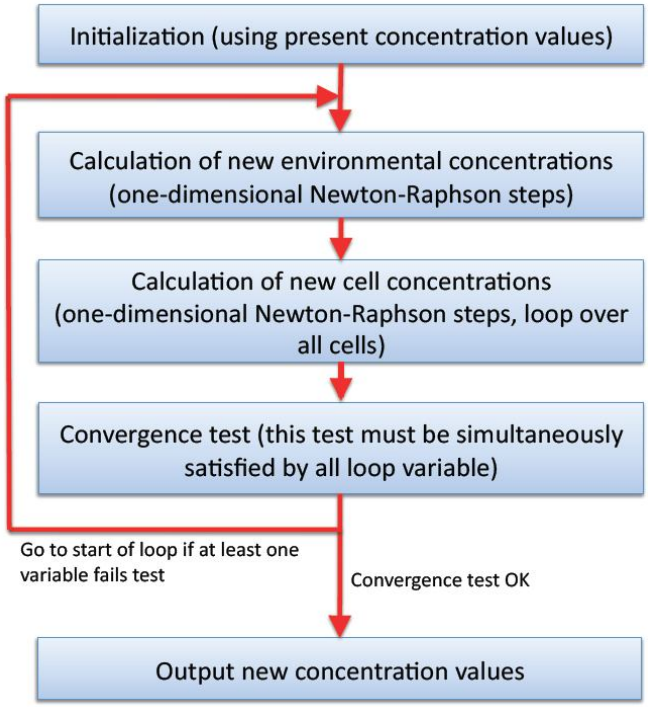
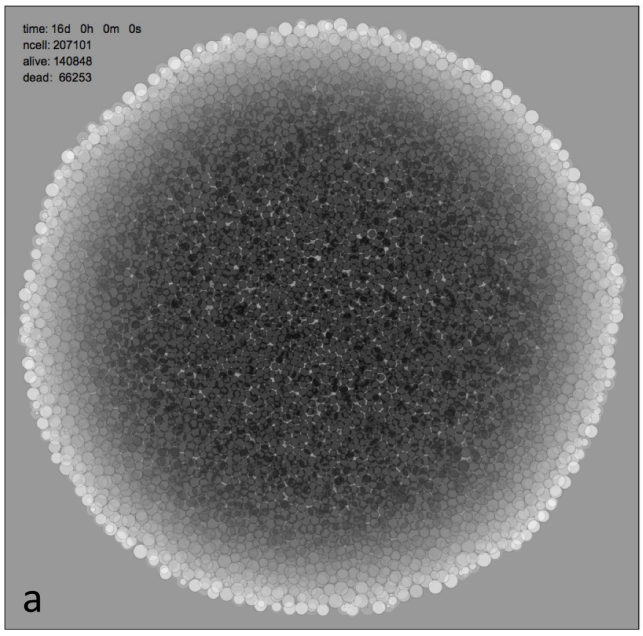


Fig. 2. Structure of loop that solves the nonlinear set of equations (2). We use one-dimensional Newton-Raphson steps (instead of taking more efficient steps that work at the full dimensionality of the variable space) because in this way the time complexity of the whole algorithm is only O(N) (where N is the number of cells); if we used the full dimensional method we would need a matrix inversion, and the time complexity would be an intractable O(N^3). Moreover, at each time step the initial concentration values are actually very close to the solution of (2), because the cluster structure changes much more slowly than diffusion, and this means that the method converges even with a extremely large number of cells.



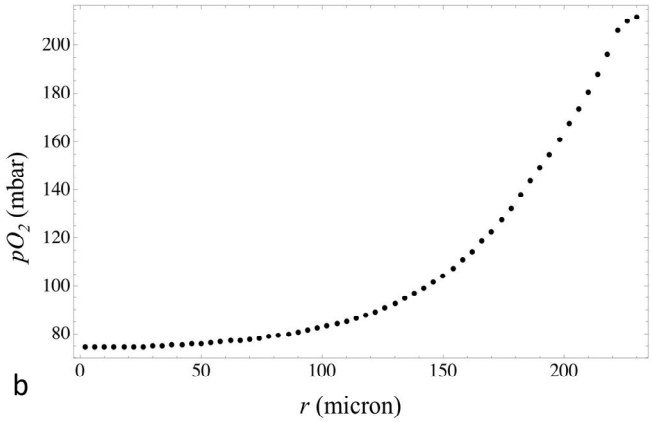


Fig 3. Oxygen concentration in a simulated spheroid. The spheroid is 16 days old (simulated time) and contains more than 200000 cells (about 140000 live cells and about 67000 dead cells). Upper panel (a): oxygen concentration in a central slice of the simulated spheroid. Oxygen concentration is mapped with different levels of gray (white = high, black = low; background is uniformly light gray to improve visibility). Lower panel (b): plot of oxygen concentration (partial pressure of dissolved oxygen, normalized to atmospheric partial pressure, in mbar), vs. the distance r from the centroid of the simulated spheroid.

REFERENCES

[1] R. Chignola and E. Milotti, "Numerical simulation of tumor spheroid dynamics." Physica A 338, 261-266 (2004).

[2] R. Chignola, E. Milotti, "A phenomenological approach to the simulation of metabolism and proliferation dynamics of large tumor cell populations." Phys. Biol. 2, 8-22 (2005).

[3] R. Chignola, A. Del Fabbro, C. Dalla Pellegrina, E. Milotti, "Ab initio phenomenological simulation of the growth of large tumor cell populations." Phys. Biol. 4, 114-133 (2007).

[4] http://www.openmp.org

[5] http://gcc.gnu.org/

[6] http://software.intel.com/en-us/intel-compilers/

[7] see, e.g., M. de Berg, O. Cheong, M. van Kreveld, and M. Overmars, *Computational geometry: algorithms and applications*. Berlin: Springer-Verlag, 2008, 3rd ed.

[8] H. Edelsbrunner and E. P. Mücke, "Three-dimensional Alpha Shapes" ACM Trans. on Graphics 13, 43-72 (1994).

[9] these calculations are performed with the help of the computational geometry library CGAL (http://www.cgal.org), which contains the implementations of many geometrical constructs, and in particular it contains methods for the fast calculation of Delaunay triangulations and alpha-shapes.

[10] E. Milotti, A. Del Fabbro, R. Chignola, "Numerical integration methods for large-scale biophysical simulations." Comput. Phys. Commun. 180, 2166-2174 (2009).

[11] R. Chignola and E. Milotti, "Tumor microenvironment in a real-life model of tumor spheroids", these Proceedings

[12] S. Walenta, J. Doetsch, W. Mueller-Klieser, L. A. Kunz-Schughart, "Metabolic imaging in multicellular spheroids of oncogene-transfected fibroblasts." J Histochem Cytochem 48, 509-22 (2000).

[13] W. Mueller-Klieser, J. P. Freyer, R. M. Sutherland, "Influence of glucose and oxygen supply conditions on the oxygenation of multicellular spheroids." Br J Cancer 53, 345-53 (1986).

[14] J. Alvarez-Pérez, P. Ballesteros, S. Cerdán, "Microscopic images of intraspheroidal pH by 1h magnetic resonance chemical shift imaging of pH sensitive indicators." MAGMA 18, 293-301 (2005).

[15] H. Acker, J. Carlsson, G. Holtermann, T. Nederman, T. Nylen "Influence of Glucose and Buffer Capacity in the Culture Medium on Growth and pH in Spheroids of Human Thyroid Carcinoma and Human Glioma Origin." Cancer Res 47, 3504-3508 (1987).

[16] R. M. Sutherland, B. Sordat, J. Bamat, H. Gabbert, B. Bourrat, et al., "Oxygenation and Differentiation in Multicellular Spheroids of Human Colon Carcinoma." Cancer Res 46, 5320-5329 (1986).

A Computational Cell Based Multi-scale Multi-model Framework for the Prediction of Cell

Tomas Bily, Vojtech Bednar, Michal Karasek, and Tomas Mikula

***Abstract*—Mathematical and computational multi-scale multi-model models of complex natural processes are being intensively studied. A natural way to describe and predict complex biological processes is to use multi-scale multi-model models. We describe a computational cell based ('middle-out') multi-scale multi-model framework for the prediction of complex biological processes and we demonstrate the usage of the framework in cancer biology.**

I. Introduction

OVER the past few decades various mathematical models have been developed in order to describe biological phenomena. The main approaches adopted by these models are continuous mathematics, discrete mathematics , hybrid, and agent based. Each modeling approach, discrete or continuous, assumes a typical spatial and temporal scale and a specific nature of the physical or biological interactions involved. Mean-field-like and local short-range interaction possibly may occur. In addition, the balance between determinism and stochasticity depends on the time and length scale on which questions are posed [1].

There has been a growing interest in a cellular based approach ('middle-out' approach [2]) for building multi-scale models. The main interest is focused on a cellular level that is gradually expanded to include lower and higher levels of bio-complexity. There are a few basic categories of multi-scale approaches used. The categories are cellular-continuum, spatially hierarchical, temporary separated approach [1]. A generalization has been achieved by the complex automata approach [3].

II. The Cell Based Multi-scale Multi-model Framework

A. The Multi-Peel Framework

The complex automata approach (CxA) uses a network of cellular automata (CA) each one operating on a predefined time and space scale that communicate through an edge defined description of interactions by utilizing a propagation-collision paradigm [3].

For a better refinement of the description of a biological process we extended the idea of CxA. To incorporate data flow and control flow between different models of processes and separate data from control we have developed a multi-peel framework. A peel represents a model of biological process or sub-process that take place in defined spatial and temporal scales, the input parameters with their domains, the state parameters/variables with their domains and a set of rules that govern the evolution of the model. We have generalized the notion of CxA by a peel graph (PG). A PG is defined by a pair of vertices and oriented edges so that vertices represent peels and oriented edges represent a description of the interaction (coupling interaction) which mostly refers to how to map the state/input parameters of a given pair of peels. The description of the interaction is divided into two parts. The data flow driven part and the control flow driven part. Edges are labeled. The label 'integrate' marks edges that represent the integration of a source peel (the source peel is being integrated into a destination peel), the label 'refine' marks edges that represent refining of the source peel (mostly backward oriented edges with the label 'integrate'.) The label 'interact' marks the rest of the edges. The labels given create a hierarchy of the integration of the models. The hierarchy can be viewed as follows: nodes having no outgoing 'integrate' edges belong to the lowest level. The next level contains nodes that have outgoing 'integrate' edges into the last building level. The execution of the model on a peel-graph is performed by executing the peels processes. Each peel has a defined time during which it can be executed or be stopped. Edges labeled as 'integrate' are handled by means of a propagation-collision process (the peel processes are executed independently and communicate with others asynchronously). Edges labeled as 'integrate' and 'refine' are handled by a *gather-update* process. The peel execution process waits until the communication among all 'integrate' edges has been completed and then continues.

Manuscript received April 20, 2010. This work was supported in part by the Istituto Nazionale di Fisica Nucleare (I.N.F.N.).

E. Milotti is with the Dipartimento di Fisica, Università di Trieste and I.N.F.N. – Sezione di Trieste, Via Valerio 2, I-34127 Trieste, Italy (phone: +39-040-558-3388; fax: +39-040-558-3350; e-mail: edoardo.milotti@ts.infn.it).

R. Chignola, is with the Dipartimento di Biotecnologie, Università di Verona and with the I.N.F.N. – Sezione di Trieste, Strada Le Grazie 15 CV1, I-37134 Verona, Italy (e-mail: roberto.chignola@univr.it).

B. Modeling tools for the description of a biological cellular processes - The Digital Biological Cell (DBC) framework

In the multi-peeling context we have built a framework for the description of cellular processes. It is a suite of tools that helps describe cellular processes and connect them together in a peeling graph [4]. The first group of tools is for the description of micro-environmental (extra, intra) processes such as diffusion and reaction kinetics regarding the concentration of chemicals, ligands and receptors. The second group of tools is for the description of cellular processes. The cell can be defined as a individual entity that consists of one or more compartments (cytosol, nucleus etc.) including compartments with a membranous character (such as cellular/nuclear membrane) and can be located in a specific environment (space). The surface of a membrane contains receptors. A cell can have a defined volume and/or shape that can evolve. The state of a cell is described by the cell cycle phase and the cell phenotype. Transition between cell cycle phases and phenotypes can be described by a graph of transitions. Vertices are pairs of phase and phenotype. Edges are labeled by conditions when transitions can happen. A condition can be applied on the current solutions of the system of equations describing an interaction network. The graph of actions describes what action (production of ligands/chemicals, binding receptors, movement, death ...) can be performed by a cell in a given state. A tool translates the description of the interaction network into a system of Ordinary Differential Equations (ODEs). The third group of tools is meant for the description of extra-cellular long range interactions such as a physical interactions in tissue (e.g. pressure propagation or convective flow). One of the tools solves the problem using the cellular Potts model (CPM [5]).

We implemented the mentioned frameworks as a modular system in C++ and Python languages that employs the SciPy library (http://www.scipy.org). The implementation of peels strictly separates the data definition and manipulation from from the functional parts. Because the definition of peels is highly structured and hierarchical and every part has a unique name in a dot name hierarchy it is possible to easily modify existing peel models thus achieving a high re-usability of code. Due to the need for a simple definition of the peel models and the PG structure a PG modeling language (PiGML) was developed. PiGML is an XML [5] like language that binds the underlaying structure into the XML hierarchy and attributes that can be extended by Python code. We have tried to follow the ideas of CellML [6] and MML [7]. Several approaches to the creation of a new peel or the changing of an implemented one, the connection of peels by edges and the definition of edge control and data flow have been developed.

III. EXAMPLES OF PG MODELS

A. Tumor evolution

Case 1) We have followed a model of avascular tumor growth described in [8] and reformulated it in the PG approach. We have determined four peels. The first peel describes the extra-cellular environment (a 3D homogenous lattice) and processes such as diffusion of chemicals and ligands (oxygen, glucose, metabolic waste, growth promoters and inhibitors). The second peel describes the cell long range behavior by CPM. The third peel represents the decision/action cellular processes (consumption, production, division, quiescence, death). The fourth peel describes the the cell cycle process. A cell is defined as one compartmental entity and the compartment is used for the communication with the extracellular environment. The extracellular peel is governing the execution of the cellular process peel and the CPM peel. Synchronization is done after 1/4 of a Monte Carlo step (MCS [8]). The time of four MCS is equal to one cell cycle. We have run a simulation for the parameter values of [8] and the results have been compared with the results of [8] to check whether the implementation was correct.

Case 2) In order to demonstrate the ease in modifying the model we have considered the model implemented in case 1 and made a new peel describing cell-cycle process according to [9]. We have added terms describing the influence of generic growth promoters and inhibitors. We have adapted the parameters of [9] and run a new series of simulations. We have observed that that both models behave similarly.

Case 3) To demonstrate an extension of the previous model regarding the cell migration potential we have adapted the EGFR gene-protein regulation network appearing in [10] so that it becomes a new peel. We have added a new term in the CPM Potts Hamiltonian to describe chemotaxis [14] and an edge into the PG that switches the chemical potential parameter from zero to non-zero depending on the phenotype. We have refined the cell compartmental description to three compartments (each one represented as a new peel): the cellular membrane, the cytosol and the nucleus. We have updated the state transitional rules to follow the rules of [10].

Case 4) A preliminary implementation of mechano-receptor induced behavior has been adapted based on the models of [11] and [12].

B. Tumor Induced Angiogenesis

In order to investigate the influence of tumor induced angiogenesis we have translated a model of [13] into the PG approach according to the description of case A.1. A

preliminary investigation of the behavior of the combined tumor and angiogenesis model has been performed.

IV. Conclusion

We have presented a computational cell-based multi-scale multi-model framework that allows to build or combine different models of biological processes. We have demonstrated the potential of the approach in the case of cancer biology modeling. Our current work focuses on the improvement of the framework and the further investigation of the interactions of different constituting models.

Acknowledgment

We would like to thank Zdenek Hedrlin for his inspiring attitude towards the work presented. His insistence on the development of new ideas and interpretations has been the most delightful experience in our research trajectory.

References

[1] M.L. Martins, S.C. Ferreira Jr., M.J. Vilela: Multiscale models for the growth of avascular tumors, Physics of Life Rev. 4 (2007) 128-156

[2] D.C.Walker and J.Southgate: The virtual cell - a candidate co-ordinator for 'middle-out' modeling of biological systems. Briefings in Bioinformatics 10 (2009), 450-461.

[3] A. Hoekstra, E. Lorenz, J.-L. Falcone, and B. Chopard: Towards a Complex Automata Framework for Multi-scale Modeling: Formalism and the Scale Separation Map, in Computational Science – ICCS 2007 (LNCS vol. 4487, Springer Verlag) 2007, pp. 922-930.

[4] T.Bily, M.Balek, V.Bednar, M.Karasek. Microscopic tumor growth module for ContraCancrum project. Proceedings of World Congress on Medical Physics and Biomedical Engineering September 7 - 12, 2009 Munich, Germany.

[5] Extensible Markup Language (XML): http://www.w3.org/XML/

[6] CellML: http://www.cellml.org/

[7] J.-L. Falcone, B. Chopard, A. Hoekstra: MML toward a Multiscale Modeling Language. Procedia Computer Science 1 (2010), 819-826

[8] Y. Jiang, J. Pjesivac-Grbovic, Ch. Cantrell, J.P.Freyer: A Multiscale Model for Avascular Tumor Growth. Biophys. J. 89 (2005), 3884-3894.

[9] T. Alarcon, H.M. Byrne, P.K. Maini. A mathematical model of the effect of hypoxia on the cell-cycle of normal and cancer cell. J. Theor. Biol. 229 (2004), 395-411.

[10] L. Zhang, Ch.A. Athale, T.S. Deisboeck: Development of a three-dimansional multiscale agent-based tumor model: Simulation gene-protein interaction profiles, cell phenotypes and multicellular patterns in brain cancer. J. Theor. Biol. 244 (2007), 96-107.

[11] I. Ramis-Conde, D. Drasdo, A.R.A. Anderson, M.A.J. Chaplain: Modeling the Influence of the E-Cadherin-beta-Catenin Pathway in Cancer Cell Invasion: A Multiscale Approach, Biophys. J. 95 (2008), 155-165.

[12] B.M.Rubenstain, L.J.Kaufman: The Role of Extracellular Matrix in Glioma Invasion: A Cellular Potts Model Approach. Biophysical J. 95 (2008), 5661-5680.

[13] Amy L. Bauer , Trachette L. Jackson and Yi Jiang: A Cell-Based Model Exhibiting Branching and Anastomosis during Tumor-Induced Angiogenesis, Biophysical Journal 92:3105-3121 (2007)

[14] M. Alber, M. Kiskowski, J. Glazier and Y. Jiang: "On Cellular Automaton Approaches to Modeling Biological Cells", in *IMA 134: Mathematical systems theory in biology, communication, and finance. Springer-Verleg, New York.* , pp. 12, 2002

A Continuum Model of Mesenchymal Cell Migration and Sprouting Angiogenesis

Florian Milde, Michael Bergdorf and Petros Koumoutsakos

***Abstract*— Collective cell motion is fundamental to processes such as organ formation, invasive tumor growth and tumor induced angiogenesis. We present a computational framework based on a diffuse interface approach for the simulation of such collective cell motion with an emphasis on mesenchymal cell migration and sprouting angiogenesis. The model explicitly considers cell-cell and cell-matrix adhesion, extracellular matrix guidance of migrating cells and chemotaxis.**

I. Introduction

CELL migration is fundamental to a number of physiological processes such as embryogenesis, organ growth, inflammation, wound healing, tumor invasion and tumor induced angiogenesis.

All these processes are characterized by a collective, directed motion of the cells in play. In many cases the migration direction is aligned with a chemical gradient, *e.g.* of nutrients or a growth factor; in this case the migration is referred to as chemotaxis.

Next to chemotactic guidance, collective cell migration is largely regulated by cell-cell adhesion and adhesion to the extracellular matrix (ECM). Cell migration in response to adhesion gradients is referred to as haptotaxis.

One particular aspect of cell migration which has attracted increased attention recently is related to the process of angiogenesis as induced by tumor cells emitting Vascular Endothelial Growth Factors (VEGF). Tumor-induced formation of new blood vessels in the context of sprouting angiogenesis is based on the orchestrated migration of clusters or cords of endothelial cells. The endothelial cells migrate through the extracellular matrix following chemotactic and haptotactic cues. For these endothelial cells, the most prominent chemoattractant is VEGF[1-2].

Existing models of sprouting angiogenesis account for chemotaxis [3-7]. The extracellular matrix is considered in [3] and [6]. In a recent work, we introduced a hybrid model of sprouting angiogenesis that explicitly considers the structure of the extracellular matrix [7]. For a detailed review on existing models of angiogenesis, please refer to [8] and references therein.

A grand majority of these models are based on cell-based representations. In a continuum formulation, one basic problem is the representation of cell-cell adhesion, as the notion of single cells does not exist.

In the present article, we focus on the continuum formulation of a computational efficient cell-cell adhesion model, and the interaction of in-silico cells with the underlying artificial extracellular matrix. Painter [9] employed a diffuse interface approach to study cell organization and the incipience of fingering patterns in the extracellular matrix. In contrast to the work presented in [9], we propose to model cell-cell adhesion local via the introduction of a proxy chemoattractant.

Manuscript received May 5, 2010.

Florian Milde is with the Chair of Computational Science, ETH, Zurich, 8092 Switzerland (e-mail: mildef@ethz.ch).

Michael Bergdorf was with the Chair of Computational Science, ETH, Zurich, 8092 Switzerland. He is now with D. E. Shaw Research, New York, NY 10036 USA (e-mail: bergdorf@inf.ethz.ch).

Petros Koumoutsakos is with the Chair of Computational Science, ETH, Zurich, 8092 Switzerland. (phone: 044-632-5258; fax: 044-632-1703; e-mail: petros@ethz.ch).

II. A Continuum Model of Cell Migration

A. Continuum Cell Representation

We choose to represent a single cell populations by a density ρ. This density evolves in time according to:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\mathbf{a}\rho) = d\Delta\rho + R(\rho), \qquad (1)$$

where $\mathbf{a}$ symbolizes the cumulative effect of cell-cell adhesion $(\mathbf{a}^{c/c})$, pressure $(\mathbf{a}^{p})$ and migration cues $(\mathbf{a}^{ecm,\phi})$. The right hand side accounts for random fluctuations on the cell population modeled at a macroscopic scale by a diffusion term with diffusion coefficient d and a reaction term $R(\rho)$ to account for proliferation and cell death. If more then one cell type is present, one density is used per type, *i.e.* $\{\rho_i\}_{i=1}^{\#CellTypes}$.

B. Cell-Cell Adhesion

Cell-cell adhesion is a fundamental biophysical mechanism. It is responsible for tissue formation, stability and breakdown. It is involved in tissue invasion and metastasis of tumor cells.

Cell adhesion to another cell or the ECM is established by specific adhesion receptors on the cell membrane. This reaction is very local, as it happens upon contact. We state a set of requirements that reflect the main characteristics of cell adhesion: (i) it is a short-range force; (ii) it will give rise to a movement of the cells towards the entity they adhere to.

Given these characteristics we can model cell adhesion as an autocrine (in the case of cell-cell adhesion), or paracrine (in the case of cell-ECM adhesion) like signal f_i acting as an adhesive force $\mathbf{a}_i^{c/c}$ on the cell population ρ_i.

$$\mathbf{a}^{c/c}{}_i = \sum_j^{\#CellTypes} \kappa_{ij} L\left(f_j, df_j\right) \nabla f_j,$$
$$\frac{\partial f_i}{\partial t} = -\mu_i f_i + \alpha_i \left(1 - \frac{f_i}{f_{i,\max}}\right) \rho_i + D_i \Delta f_i, \quad (2)$$

μ_i, α_i and D_i are the decay, release and diffusion parameters of the adhesion signal of cell population ρ_i. $f_{i,\max}$ denotes the threshold value for the release of f_i. The factor κ_{ij} describes the heterotypic $(i \neq j)$ and homotypic $(i = j)$ adhesion strength. $L\left(f_i, df_i\right) = df\left(\max\left(df_i, |\nabla f_i|\right)\right)^{-1}$ is a cutoff function that keeps the magnitude of the gradient bounded by df_i.

C. Close-packing Density

We incorporate repulsive effects that might limit the local cell density by adding the following pressure-like term to the velocity:

$$\mathbf{a}^p = -\kappa_p H\left(\rho - \bar{\rho}\right) \nabla\rho \left|\nabla\rho\right|^{-1}, \quad (3)$$

where $\rho \equiv \sum_i \rho_i$, κ_p is a constant that determines the cell population response to pressure, $\bar{\rho}$ is the cell close-packing density and H is the Heaviside function.

D. The Extracellular Matrix

Continuum simulations of cellular motion rarely explicitly consider the effects of the ECM on migration. These effects are however crucial, as the ECM serves as a scaffolding with adhesive sites which the cells can use to exert forces and propel themselves. We propose to model the ECM as a collection of bundles of adhesive fibers, which facilitate but also bias migration. The matrix is constructed by distributing N_f fiber bundles of predefined length and thickness randomly throughout the computational domain.

These fibers are then discretized onto the ECM grid e using Bresenham's line rasterization algorithm. In order to get a differentiable field we filter e with a second-order B-spline kernel N_{filter}-times.

E. Chemotaxis and Adhesion Inside the ECM

Our formulation of chemotaxis is based on the most simple approach, where the migrating cells follow the concentration gradient of a chemoattractant ϕ. The model extension presented here to account for cell-ECM adhesion is a formulation of the following assumptions: in order to maximize its migration velocity, a cell will crawl along fibers, if these fibers are not transverse to the chemotactic cue $(\nabla\phi)$. If there are no fibers in its environment, *i.e.* $e = 0$, then a cell will not be able to migrate efficiently $(e_o << 1)$, if the fiber density is too high $(e \approx e_\infty)$, then cells have to degrade the matrix before they are able to migrate. These assertions are represented by:

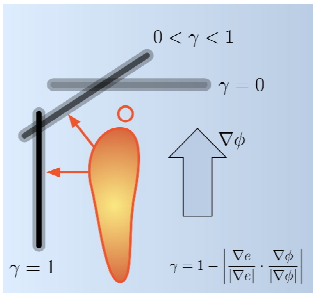


Fig. 1. Conceptual sketch of the model of cell-matrix adhesion. A cell will move "onto" a fiber if the fiber direction is not transversal to the chemotactic gradient, *i.e.* the gradient of adhesion is not aligned with the chemtactic direction.

$$\mathbf{a}^{ecm,\phi} = \left[\left(1 - \left|\frac{\nabla e}{|\nabla e|} \cdot \frac{\nabla\phi}{|\nabla\phi|}\right|\right) \nabla e + \nabla\phi\right] \left(e + e_o\right)\left(e_\infty - e\right). \quad (4)$$

The parameter e_o defines the minimal migrative response in the total absence of an ECM, e_∞ defines the matrix density threshold that completely blocks migration.

III. Results

We have applied the method to study the influence of the ECM density on the process of angiogenesis (Figure 2). Endothelial cells lining an existing vasculature are initialized in a wavy front to the left side of the domain.

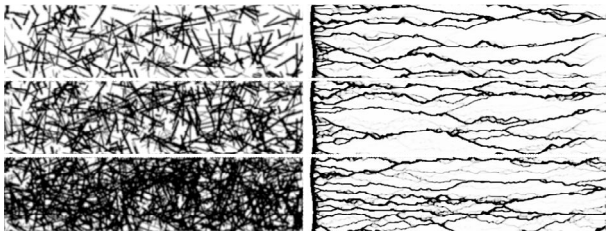


Fig. 2. Effect of increasing the matrix density (top to bottom) on the process of angiogenesis. Left: Extracellular matrix of density 31%, 51% and 75%. Right: Vascular network grown from left to right.

We assume a linear chemotactic signal increasing towards the right side of the domain. For an increase in the number of distributed fibers (Figure 2, left panel) we observe a strong effect on the morphology of the final network structure of endothelial cell density (Figure 2, right panel). An increase in the matrix density results in an increase in the number of observed branches in combination with a decreased vessel segment length (data not shown).

REFERENCES

[1] N. Ferrara H.-P. Gerber, J. LeCouter, *Nature Medicine* vol. 9, pp. 669-676, 2003.

[2] J. Folkman, *Annual Review of Medicine*, vol. 57 pp. 1-18, 2006.

[3] A. J. Bauer, T. J. Jackson, Y. Jiang, Biophysical Journal, vol. 92, pp. 3105-3121, 2007.

[4] A. R. A. Anderson, M. A. J. Chaplain, *Bulletin for Mathematical Biology*, vol. 60, pp. 1998.

[5] M. A. Chaplain, *Journal of Neuro-Oncology*, vol. 50, pp. 37-51, 2000.

[6] S. Sun, M. F. Wheeler, M. Obeyesekere, J. Patrick, *Bulletin of Mathematical Biology*, vol. 67, pp. 313-337, 2005.

[7] F. Milde, M. Bergdorf, P. Koumoutsakos, *Biophysical Journal*, vol. 95, pp. 3146-3160, 2010.

[8] J. S. Lowengrub *et al.*, *Nonlinearity*, vol. 23, pp. R1-R91, 2010

[9] K. Painter, *Journal of Mathematical Biology*, vol. 58, pp. 511-543, 2009.

IV. Cellular and Higher Level Based Modeling: Predominantly *Top-Down* Approaches

In Silico Oncology: A Hypermatrix–Operator Formulation of a Top-Down Multiscale Simulation Model of Tumor Response to Treatment. The *Oncosimulator* Concept

Georgios S. Stamatakos

***Abstract*— A hypermatrix-operator formulation of a multiscale top-down tumor dynamics modeling method primarily based on the consideration and manipulation of discrete biological entities and events is outlined. The method is clearly clinically oriented. One of its major goals is to support patient individualized treatment optimization through experimentation *in silico* (=on the computer). Therefore, modeling of the treatment response of clinical tumors lies at the epicenter of the approach. Macroscopic data, including i.a. anatomic and metabolic tomographic images of the tumor, provide the Virtual Physiological Human (VPH) framework for the integration of data and mechanisms pertaining to lower and lower biocomplexity levels such as clinically approved cellular and molecular biomarkers. The method also provides a powerful framework for the investigation of multiscale tumor biology in the generic investigational context. The *Oncosimulator*, a multiscale basic science and biomedical engineering concept and construct tightly associated with the method is also sketched. The current status of the simulation executions, clinical adaptation and validation of the corresponding models is briefly outlined. The paper concludes with a brief discussion of several aspects of the approach and future perspectives.**

I. Introduction

Most cancer modeling techniques developed up to now adopt the straightforward *bottom-up* approach focusing on the better understanding and quantification of rather microscopic tumor dynamics mechanisms and the investigation of crucial biological entity interdependences including i.a. tumor response to treatment in the generic investigational context. To this end several combinations of mathematical concepts, entities and techniques have been developed and/or recruited and appropriately adapted. They include inter alia population dynamics models, cellular automata and hybrid techniques, agent based techniques, diffusion related continuous and finite mathematics treatments etc. In addition, a number of large clinical tumor models focusing mainly on invasion and tumor growth morphology rather than on tumor response to concrete therapeutic schemes as administered in the clinical setting have appeared. Finite difference and finite element based solutions of the diffusion and classical mechanics equations constitute the core working tools of the corresponding techniques.

Nevertheless, a number of concrete and pragmatic clinical questions of importance such as the prediction of tumor response to a given treatment scheme or the selection of the optimal candidate scheme for a given patient cannot be dealt with neither by the bottom-up approach nor by the morphologically oriented large tumor growth models in a direct and efficient way. A promising modeling method designed with the primary aim of answering such questions is the *top-down* method developed by the *In Silico* Oncology Group (www.in-silico-oncology.iccs.ntua.gr) [1-8].

Macroscopic data, including i.a. anatomic and metabolic tomographic images of the tumor and the broader anatomic region of interest, provide the framework for the integration of available and clinically trustable biological information pertaining to lower and lower biocomplexity levels such as clinically approved histological and molecular markers. However, the method does also provide a powerful framework for the investigation of multiscale tumor biology in the generic investigational context.

From the mathematical point of view the top-down simulation method outlined is primarily a discrete mathematics method, although continuous mathematics (continuous functions, differerential equations) are used in order to tackle specific aspects of the models such as pharmacokinetics and cell survival probabilities based on

Manuscript received May 5, 2010. (This work has been supported in part by the European Commission under the projects "ACGT: Advancing Clinicogenomic Trials on Cancer" (FP6-2005-IST-026996) and ContraCancrum: Clinically Oriented Translational Cancer Multilevel Modeling" (FP7-ICT-2007-2- 223979)).

G. S. Stamatakos is with the Institute of Communication and Computer Systems, National Technical University of Athens, 9 Iroon Polytechniou, GR 157 80, Zografos, Greece (phone: +30 210772 2287; fax: +30 2107733557; e-mail: gestam@central.ntua.gr).

pharmacodynamical and radiobiological models. Adoption of the discrete approach as the core mathematical strategy of the method has been dictated by the obvious fact that from the cancer treatment perspective it is the discrete i.e. the integer number of the usually few biological cells surviving treatment and their discrete mitotic potential categorization (stem cells, progenitor cells of various mitotic potential levels and differentiated cells) that really matter. These discrete entities and quantities in conjunction with their complex interdependences and stochastic behavior may give rise to tumor relapse or to ensure tumor control over a given time interval following completion of the treatment course. Cell cycle phases have a clearly discrete character too. Moreover, the properties of the different cell phases may vary immensely from the clinical significance perspective. A classical example is the lack of effect of cell cycle specific drugs on living tumor cells residing in the resting G0 phase. The *top-down* method under consideration constitutes the core simulation strategy of the Oncosimulator [5].

II. The Oncosimulator Concept

The Oncosimulator [5] is a concept of multilevel integrative cancer biology, a complex algorithmic construct, a biomedical engineering system and (eventually in the future) a clinical tool which 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.

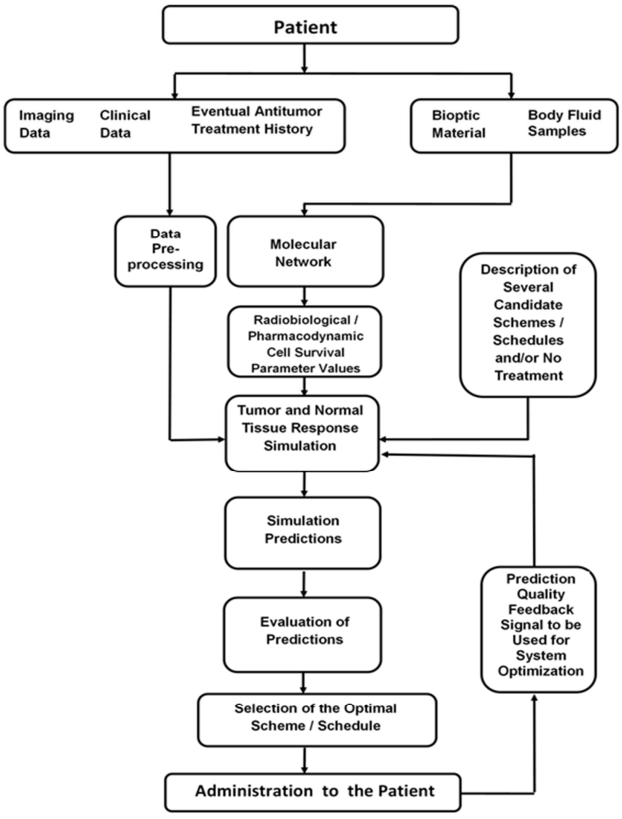


Fig.1 The Oncosimulator: a gross worksflow diagram

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. Fig. 1 shows a synoptic diagram of its envisaged functioning.

III. The Basics of the Top-Down Method

The anatomic region of interest is discretized by a virtual cubic mesh of which the elementary cube is termed geometrical cell. A *hypermatrix* i.e. a mathematical matrix of (matrices of (matrices…of (matrices or vectors or scalars))) corresponding to the anatomic region of interest is subsequently defined [4,5]. The latter describes explicitly or implicitly the local biological, physical and chemical dynamics of the region. The following (sets of) parameters are used in order to identify a cluster of biological cells belonging to a given equivalence class within a geometrical cell of the mesh at a given time point: I. the spatial coordinates of the discrete points of the discretization mesh with spatial indices *i, j, k* respectively. It is noted that each discrete spatial point lies at the center of a geometrical cell of the discretization mesh. II. the temporal coordinate of the discrete time point with temporal index *l*. III. the mitotic potential category (i.e. stem or progenitor or terminally differentiated) of the biological cells with mitotic potential category index *m*. IV. the cell phase (within or out of the cell cycle) of the biological cells with cell phase index *n*. The following phases are considered: {G_1, S, G_2, M, G_0, A, N, D}, where G_1 denotes the G_1 cell cycle phase; S denotes the DNA synthesis phase; G_2 denotes the G_2 cell cycle phase; M denotes mitosis; G_0 denotes the quiescent (dormant) G_0 phase; A denotes the apoptotic phase; N denotes the necrotic phase and D denotes the remnants of dead cells.

For the biological cells belonging to a given mitotic potential category AND residing in a given cell phase AND being accommodated within the geometrical cell whose center lies at a given spatial point AND being considered at a given time point; in other words for the biological cells clustered in the same equivalence class denoted by the index combination *ijklmn*, the following state parameters are provided: i. local oxygen and nutrient provision level, ii. number of biological cells, iii. average time spent by the biological cells in the given phase, iv. number of biological cells hit by treatment, v. number of biological cells not hit by treatment.

The initial constitution of the tumor i.e. its biological,

physical and chemical state has to be estimated based on the available medical data through the application of pertinent algorithms. This state corresponds to the instant just before the start of the treatment course to be simulated. The entire simulation can be viewed as the periodic and sequential application of a number of sets of algorithms (*operators*) on the hypermatrix of the anatomic region of interest. The application of the operators on the hypermatrix of the anatomic region of interest takes place in the following order: A. Time updating i.e. increasing time by a time unit (e.g. 1h), B. Estimation of the local oxygen and nutrient provision level. C. Estimation of the effect of treatment (therapy) referring mainly to cell hitting by treatment, cell killing and cell survival. Available molecular and/or histological information is integrated primarily at this point. D. Application of cell cycling, possibly perturbed by treatment. Transition between mitotic potential cell categories such as transition of the offspring of a terminally divided progenitor cell into the terminally differentiated cell category is also tackled by this algorithm set. E. Differential expansion or shrinkage or more generally geometry and mechanics handling. F. Updating the local oxygen and nutrient provision level following application of the rest of algorithm sets at each time step. It is noted that stochastic perturbations about the mean values of several model parameters are considered (hybridization with the Monte Carlo technique).

IV. Simulations - Clinical Adaptation – Validation

The simulation codes for a number of tumor types (glioblastoma multiforme, breast cancer, nephroblastoma) treated according to various clinical trial schedules have been executed in order to check and optimize the corresponding models and explore their behavior. To this end series of anonymized multiscale data stemming from clinical trials such as the nephroblastoma SIOP 2001/GPOH trial have been exploited [3]. Agreement between the real and the simulated time course of the tumor volume before, during and after chemotherapeutic treatment of a clinical tumor has been used as the primary clinical adaptation and validation criterion. Adaptation of the model to imaging reality for a small number of multiscale clinical trial data sets using plausible parameter values has been achieved. Further information on the ongoing clinical adaptation and validation process is available in [3].

V. Conclusions

Both the *top-down* multiscale cancer simulation method briefly outlined above and the *Oncosimulator* concept have been designed so as to be readily optimizable, extensible and adaptable to changing clinical, biological, and research environments. They constitute the simulation core of both the ACGT and the ContraCancrum European Commission funded research projects (see Acknowledgements). An ongoing large scale sensitivity study of the corresponding models as well as an ongoing clinical optimization and validation process are expected to shape and support the envisaged clinical translation of both entities.

References

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

[2] Dionysiou, D.D., G.S. Stamatakos, D. Gintides, N. Uzunoglu K. Kyriaki. 2008. Critical parameters determining standard radiotherapy treatment outcome for glioblastoma multiforme: a computer simulation. Open Biomed Eng J 2: 43-51.

[3] Graf, N., A. Hoppe , E. Georgiadi, R. Belleman, C. Desmedt, D. Dionysiou, M. Erdt , J. Jacques, E. Kolokotroni, A. Lunzer, M. Tsiknakis and G. Stamatakos. 2009. "*In silico* oncology" for clinical decision making in the context of nephroblastoma. Klin Paediatr 221: 141-149.

[4] Kolokotroni E.A.,, G. S. Stamatakos, D. D. Dionysiou, E. Ch. Georgiadi, Ch. Desmedt, N. M. Graf, "Translating Multiscale Cancer Models into Clinical Trials: Simulating Breast Cancer Tumor Dynamics within the Framework of the "Trial of Principle" Clinical Trial and the ACGT Project.," Proc. 8th IEEE International Conference on Bioinformatics and Bioengineering (BIBE 2008), Athens, Greece, 8-10 Oct. 2008. IEEE Catalog Number: CFP08266, ISBN: 978-1-4244-2845-8, Library of Congress: 2008907441, Paper No. BE-2.1.1, length: 8 pages (in electronic format). 2008

[5] Georgiadi E. Ch., G. S. Stamatakos, N. M. Graf, E. A. Kolokotroni, D. D. Dionysiou, A. Hoppe, N. K. Uzunoglu, "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," Proc. 8th IEEE International Conference on Bioinformatics and Bioengineering (BIBE 2008), Athens, Greece, 8-10 Oct. 2008. IEEE Catalog Number: CFP08266, ISBN: 978-1-4244-2845-8, Library of Congress: 2008907441, Paper No. BE-2.1.2, length: 8 pages (in electronic format). 2008

[6] Giatili S.G., Georgios. S. Stamatakos, Dimitra D. Dionysiou, Eleni A. Kolokotroni, Eleni Ch. Georgiadi, Geometrical and Mechanical Aspects of Tumor Growth and Response to Chemotherapeutic Schemes in the Context of the ACGT Oncosimulator," Proceedings of the 3rd International Advanced Research Workshop on In Silico Oncology, Istanbul, Turkey, Sept. 23-24, 2008. Edited by G. Stamatakos and D. Dionysiou pp.35-37, 2008

[7] Dionysiou, DD., G.S. Stamatakos, D. Gintides, N. Uzunoglu, K. Kyriaki, 2008, "Critical Parameters Determining Standard Radiotherapy Treatment Outcome for Glioblastoma Multiforme: A Computer Simulation". The Open Biomedical Engineering Journal 2, 43-51, 2008

[8] Stamatakos G. and D. Dionysiou. 2009. Introduction of hypermatrix and operator notation into a discrete mathematics simulation model of malignant tumor response to therapeutic schemes in vivo. Some operator properties. Cancer Informatics 2009:7 239–251.

[9] Stamatakos G. 2010. *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. Chapman & Hall/CRC, Boca Raton, Florida, USA, 2010 (in press) ISBN: 9781439814406.

The ISOG, NTUA Tumor Response to Treatment Discrete Simulation Models: a Review of Basic Concepts and Algorithms

Dimitra D. Dionysiou

***Abstract*—The aim of this paper is to present the basics of the discrete event/discrete entity based modeling approach developed by the *In Silico* Oncology Group, National Technical University of Athens, for solid tumor response to treatment simulations. The most prominent features and key algorithms that form the core foundation of the models are outlined.**

I. Introduction

SCIENTIFIC knowledge pertaining to the phenomenon of cancer at all levels of biocomplexity is increasing at an unprecedented rate. Multilevel integration of such information is expected to provide new powerful tools for both gaining insight into the fundamentals of the natural phenomenon of cancer and for patient-specific treatment modeling [1]. Within this context the *In Silico* Oncology Group (ISOG), National Technical University of Athens (NTUA), has formulated and developed a top-down multiscale tumor dynamics modeling method for tumor response to treatment based on the consideration and manipulation of discrete biological entities and events. The fundamentals of this approach are presented in the following sections.

II. The Basic General Features of the Approach

- *Multiscale "top-down" approach:* The top-down approach is based on the so-called "summarize and jump" strategy [2]; the method starts from the macroscopic imaging data (a high biocomplexity level) and proceeds towards lower biocomplexity levels. When there is a need for an upwards movement in the biocomplexity scales, a summary of the available information pertaining to the previous lower level is used.

Manuscript received May 5, 2010. This work was supported in part by the European Commission under the projects "ACGT: Advancing Clinicogenomic Trials on Cancer" (FP6-2005-IST-026996), and "ContraCancrum : Clinically Oriented Translational Cancer Multilevel Modeling" (FP7-ICT-2007-2-223979.)

D.D.Dionysiou is with the *In Silico* Oncology Group, Institute of Communications and Computer Systems, National Technical University of Athens, Iroon Polytechniou 9, Zografos, GR-157-80, Athens, Greece. (phone: +30-210-772-2287; fax: +30-210-772-3557; e-mail: dimdio@esd.ece.ntua.gr).

- *Discrete character:* The ISOG top-down modeling method is classifiable primarily as a discrete mathematics method (involving discrete entities and discrete events). However, certain aspects of the simulation problem, such as drug pharmacokinetics and cell survival probabilities described by pharmacodynamics and/or radiobiology can be treated with continuous mathematics.
- *Clinical orientation*: Major scientific challenge for the ISOG modeling efforts is the eventual translation of its detailed multiscale cancer models to clinical practice in the context of patient individualized treatment optimization. The models have been designed so as to respond to real clinical questions concerning the optimization of treatment strategies of individual patients and are continuously refined in the context of clinical trials [3],[4].

All models support the use of actual imaging, histopathologic, molecular and treatment data available for each particular clinical case considered. While clearly clinically oriented, they retain their potential to be used as exploratory tools in basic research. As more and more sets of medical data are exploited the reliability of the model "tuning" is expected to increase and patient individualized modeling to be strengthened.

III. Basic Constructs and Algorithms

A. "Geometrical Cells" and "Equivalence Classes"

A discretizing mesh covers the area of interest. The elementary cubic volume of the mesh is called "Geometrical Cell" (GC). The fundamental algorithmic construct of the simulation approach is a system of quantizing cell clusters within each GC of the mesh. More specifically, each GC belonging to the tumor comprises the following "Equivalence Classes" (also termed cell categories): stem cells (assumed to possess unlimited proliferative potential), cells with Limited MItotic Potential (LIMP cells), terminally differentiated cells, necrotic cells, and apoptotic cells.

Stem or LIMP cells can be proliferating or dormant (G0) (due to inadequate oxygen and/or nutrient supply). Proliferating stem or LIMP cells are further distributed

into subclasses corresponding to the cell cycle phase in which they reside: G1 (Gap 1 phase), S (DNA synthesis phase), G2 (Gap 2 phase), M (Mitosis).

B. Initialization issues

A specialized clinician delineates structures of interest on the available tomographic images of the tumor. The information derived from adequately processed imaging data directs the division of the tumor region into metabolic sub-regions, such as necrotic and well-vascularized areas of the tumor. Several model parameters can be assigned initial values based on the spatial distribution of the metabolic activity of the tumor in a clinical case-specific manner. In case that actual imaging data are unavailable, simplifying but biologically and clinically reasonable assumptions are made (e.g. consideration of macroscopically homogeneous spherical breast tumors, or triaxial ellipsoidal tumors for the nephroblastoma case). Histological data such as differentiation grade, as well as other tumor specific data, can be incorporated for a further refinement of the values attributed to the model input parameters.

The initial distribution of the proliferating cells throughout the cell cycle phases is assumed analogous to the corresponding cell cycle phases durations. Biological cells belonging to the same cell category and cell cycle phase within a given GC are assumed to be synchronized. Cells belonging to different GCs or to different categories and/or cell cycle phases within the same GC are not synchronized. Cell category transition rates are assumed constant throughout a simulation. Such an approximation is considered applicable for a relatively short time interval compared to the tumor lifetime (for example during the chemotherapeutic treatment course), and the constant values are assumed to reflect the means of the actual cell category transition rates over the interval.

Detailed studies have revealed that in order to avoid abnormal growth behaviors at the beginning of the simulation, the relative population (expressed as fraction of the total tumor cell population) of each equivalence class and its equivalence subclasses used for the population initialization of the occupied GCs should be adjusted to the values of the cell category and cell phase transition rates [5], [6]. In addition, a special condition is applied so as to eliminate cases of self-diminishing tumors [6]. Proper tumor initialization excludes the possibility of a latent artificial tumor behavior which could interfere with the interpretation of the simulation results.

C. Prominent algorithms

During a simulation the geometrical mesh is scanned and at each time step the updated state of a given GC is determined on the basis of a number of algorithms describing the spatio-temporal activity of the tumor cells. Each complete virtual scan can be viewed as consisting of two mesh scans: the first one aiming at updating the state of each GC according to the adopted cytokinetic model and the second one mainly aiming at simulating tumor shrinkage or expansion while preserving a roughly uniform cell density throughout the tumor volume.

Several progressively improved versions of specially designed stochastic cellular automata, such as the one depicted in Fig.1, describe tumor cell kinetics; this is a rather generic cytokinetic model that can be adapted for specific tumor data and treatment schemes by adequately adjusting the corresponding parameters (e.g. transition probabilities, cell cycle durations etc.) The most prominent biological phenomena explicitly described and incorporated into the latest ISOG cytokinetic models are:

- Cycling of proliferating cells through the successive phases of the cell cycle.
- Symmetric and asymmetric stem cell division.
- Proliferation and terminal differentiation of LIMP cells.
- Spontaneous apoptosis.
- Transition to a dormant (G0) phase due to inadequate supply with oxygen and nutrients.
- Local reoxygenation and nutrient provision restoration.
- Cell death through necrosis due to prolonged oxygen and nutrients deprivation.
- Chemotherapy-induced cell death. Cell cycle specific, cell cycle non specific, cell cycle phase specific and cell cycle non phase specific drugs can be readily simulated (see next section).
- Radiotherapy-induced cell death (see next section)

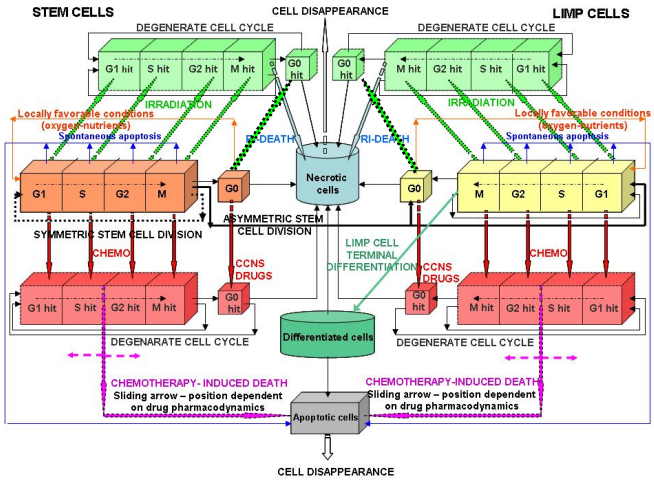


Fig. 1. Generic tumor cell cytokinetic model.. LIMP cell: Limited Mitotic Potential cell ,G1: Gap 1 phase, S: DNA synthesis phase, G2: Gap 2 phase,

M: mitosis, G0: dormant phase, hit: cells hit by drug/irradiation, CCNS drug: Cell Cycle Non-Specific Drug, RI-death: Radiation Induced death.

Several algorithms have been developed so as to simulate tumor expansion and shrinkage [7]-[9]. A typical Number of Biological Cells (NBC) is placed at all GCs during initialization. However, during a simulation the total cell population of a GC may fluctuate between a minimum, NBC_{min}, (e.g. 0.9*NBC) and a maximum, NBC_{max}, (e.g. 1.1*NBC) value.

At each time point during the second scan of the mesh, if the number of tumor cells contained within a given GC becomes less than NBC_{min}, then a procedure that attempts to "unload" the remaining few cells in the neighboring GCs takes place (26-cell-neighborhoods considered). The unloaded cells are preferentially placed into the neighboring GCs having the maximum available free space. If the given GC becomes empty, it is "removed" from the tumor. A shift of the contents of a series of GCs, intended to fill the "vacuum", leads to tumor shrinkage. This may happen after a number of cells have been killed by chemotherapy and/or irradiation.

On the other hand, if the number of cells within a given GC exceeds NBC_{max}, then a similar procedure attempting to unload the "redundant" cells in the surrounding GCs takes place. If the unloading procedure fails to reduce the number of cells to less than NBC_{max}, then a new GC "emerges". Its position relative to the "mother" GC is determined using a random number generator. Again, an appropriate shifting of the contents of a series of adjacent GCs leads to tumor expansion. The cells in the "newborn" GC are distributed in the various phase classes according to the distribution in the "mother" GC.

As far as the GC content shifting algorithms are concerned, several versions have been developed [7],[9]. More specifically:

- Version A: random selection of a shifting direction among the six possible directions (Cartesian coordinate system centered at the current GC, each semi-axis defines a shifting direction).
- Version B: detection of the outermost non-empty GC along each one of the six possible directions and counting of its "6-Neighbor" GCs belonging to the Tumor (NGCT). In the case of shrinkage, the direction corresponding to the maximum NGCT is selected. A similar, though inverse, rule may be applied in the case of tumor expansion.
- Version C: Shifting along a "line" of random direction. This algorithm is based on the generation of random points on the surface of a hypothetical sphere centered on the GC under consideration. The shifting of the GCs takes place along the "line" connecting the GC under consideration and the selected random point.
- Version D: Shifting along a "line" of random direction, coupled with counting of NGCT. This is a combination of Version C and Version B. For tumor shrinkage, the line corresponding to the maximum NGCT is selected among a number of lines of random direction.
- Version E: In case of tumor shrinkage, GC content shifting proceeds towards the tumor's center of mass, along the line connecting the tumor's center of mass with the GC to be removed from the mesh.

Thorough studies have revealed that the addition of these morphological rules improves the behavior of the system, leading to tumor shrinkage and expansion conformal to the initial shape of the tumor, on the assumption that the mechanical properties of the surrounding normal tissues are uniform. The need for the formulation of these rules has arisen from the inspection of the macroscopic results of the simulations. The goal is to avoid a premature extensive fragmentation, which is usually incompatible with clinical experience, and achieve conformal shrinkage/expansion, which is the general trend for most solid tumors. Moreover, shifting along the axes of the Cartesian coordinate system (versions A and B) produces artificial results, which become particularly obvious when considering the case of untreated spherical tumor growth; shifting along lines of random direction solves this artificial evolution problem.

The examination of the behavior of the various shifting algorithms suggests that counting of NGCT preserves the tumor connectivity. If conformal shrinking during treatment is a prerequisite, then the most advantageous algorithms are certain combinations of versions C, D and E for the expansion and shrinkage algorithms. Versions A-C which result in extensive tumor fragmentation, may be combined with an algorithm that detects GCs isolated from the main tumor mass and moves their content towards the center of the main cohesive tumor mass in a stepwise manner [8].

As far as the handling of internal necrotic regions is concerned, two approaches have been adopted, depending on the features of each simulation study (tumor type, time frame considered etc.): a) consideration of fixed boundaries between the necrotic and the proliferating region, if the simulation time frame is assumed too narrow for macroscopically apparent changes in metabolic regions to occur [9] and b) inward or outward movement of the boundary between the necrotic and the proliferating region for a shrinking or expanding, respectively, tumor [9],[10].

IV. TREATMENT SIMULATION

A. Chemotherapy

In general, for a given drug administration schedule the

concentration of the drug and its metabolites within the tumor is calculated as a function of time based on pharamacokinetic principles. Cell fates after treatment may be predicted based on the considered drug(s)' pharmacodynamics.

Currently used ISOG cytokinetic models (e.g. Fig.1) permit consideration of cell cycle–specific (CCS) or cell cycle-nonspecific (CCNS) agents. Lethally hit cycling tumor cells enter a rudimentary cell cycle that leads to apoptotic death, whereas lethally hit dormant (G0) cells enter the G0hit phase. For cell cycle phase-specific drugs (CCPS), the simulation of distinct cytotoxic mechanisms of drugs can be implemented by adjusting the relative weights of the "paths" that lead cells of various cell cycle phases to the rudimentary cell cycle, as well as by defining the position of the "exit" arrow(s) (denoting cell death) of the hit cells into the apoptotic phase.

At time instants corresponding to drug administration, the numbers of proliferating and dormant cells (stem or LIMP) hit by the drug are computed by using the cell kill ratio (CKR) parameter, defined as the percentage of cells hit by the drug after each administration.

As far as the value of the CKR parameter is concerned, two methodologies have been used depending on the nature and the goals of each particular simulation study. The "forward method", uses a priori calculated values of CKR based on values of pharmacokinetic quantities of interest and pharmacodynamic data derived from pertinent literature (e.g. experimental FDA data from cytotoxicity studies); if necessary special pharmacokinetics software (e.g. SAAM, http://depts.washington.edu/saam2) is exploited. Processed molecular data (e.g. biopsy material and/or blood) can be used to perturb the radiobiological or pharmacodynamic cell kill parameters about their population-based mean values. In the "inverse method", the value of the CKR is suggested by the clinical data and the simulation itself; it is the value that -after having selected the values of the remaining model parameters- results in good agreement between the evolution of the simulated tumor and that of the real tumor according to the clinical data (the "apparent" CKR). The "apparent CKR" for each particular clinical case, can be thought of as summarizing important genetic determinants influencing the tumor's response to therapy [6].

When combinations of drugs are administered, the assumption of an additive cell kill effect has been adopted so far (e.g additive drug effect of vincristine and actinomycin-D for the case of Wilm's tumor preoperative chemotherapy [11]).

B. Radiotherapy

Radiotherapy-induced cell death is computed based on the widely used Linear Quadratic (LQ) Model [12]. The fraction of surviving tumor cells after a uniform radiation dose D (in Gy) is given by:

$$S(D) = \exp\left[-\left(\alpha D + \beta D^2\right)\right] \quad (1)$$

where α (Gy^{-1}) and β (Gy^{-2}), termed radiosensitivity parameters, characterize the initial slope and the curvature, respectively, of the survival curve. Cell radiosensitivity varies considerably throughout the cell cycle. The S phase is regarded as the most resistant. Cells in any proliferating phase are more radiosensitive than hypoxic cells. Lethally damaged cells enter a rudimentary cell cycle and are assumed to undergo two mitotic divisions prior to their death and disappearance from the tumor [12].

Finally, the development of simulation modules of tumor response to concurrent or sequential chemotherapy and radiotherapy is underway. A critical point is to define whether the considered modalities are interactive (one modality modifying the response of the other) or non-interactive (each modality exerting its own independent effect); this requires special consideration of each particular tumor type, chemotherapeutic agent(s) and treatment schedules in a case-specific manner.

V. CONCLUSION

Ultimate goal of the ISOG discrete modeling approach, apart from contributing to an improved understanding of cancer, is the development of models that have the potential to be translated into the clinical environment, so as to serve as patient individualized treatment optimization tools, following a strict validation procedure. The models are designed and continuously refined in the context of clinical trials. In parallel, constant improvement of the algorithmic description of the involved biological mechanisms is of particular importance, in order to keep pace with the ever-accumulating scientific knowledge. The discrete and modular character of the models facilitates this challenging refinement process.

REFERENCES

[1] G. Stamatakos, "Spotlight on Cancer Informatics," *Cancer Inform*, vol. 2, pp. 83-86, 2006.

[2] G. Stamatakos and D. Dionysiou, "Introduction of hypermatrix and operator notation into a discrete mathematics simulation model of malignant tumor response to therapeutic schemes in vivo. Some operator properties," *Cancer Inform*, vol. 7, pp. 239–251, 2009.

[3] N. Graf, A. Hoppe, E. Georgiadi, R. Belleman, C. Desmedt, D. Dionysiou et al, "In Silico Oncology for Clinical Decision Making in the Context of Nephroblastoma," *Klinische Pediatrie* vol. 221, pp. 141-149, 2009.

[4] K. Marias, V. Sakkalis, A. Roniotis, C. Farmaki, G. Stamatakos, D. Dionysiou et al, "Clinically Oriented Translational Cancer Multilevel Modeling: The ContraCancrum project.," O. Doessell and W.C. Schlegel (Eds.): WC2009, *IFMBE Proceedings* 25/IV, pp. 2124-2127, 2009.

[5] E. A. Kolokotroni, G. S. Stamatakos, D. D. Dionysiou, E. Ch. Georgiadi, Ch. Desmedt, N. M. Graf, "Translating Multiscale Cancer Models into Clinical Trials: Simulating Breast Cancer Tumor Dynamics within the Framework of the "Trial of Principle" Clinical Trial and the ACGT Project.," in Proc. *BIBE 2008*, Athens, Greece, 8-10 Oct. 2008. (CFP08266, ISBN: 978-1-4244-2845-8, Library of Congress: 2008907441, Paper No. BE-2.1.1)

[6] G. S. Stamatakos, E. A. Kolokotroni,, D. D. Dionysiou, E. Ch. Georgiadi, C. Desmedt, "An advanced discrete state – discrete event multiscale simulation model of the response of a solid tumor to chemotherapy. Mimicking a clinical study," *J Theor Biol*, vol. 266 (1), pp. 124-139, 2010.

[7] D. D. Dionysiou, G. S. Stamatakos, N. K. Uzunoglu, K. S. Nikita, "A computer simulation of in vivo tumor growth and response to radiotherapy: new algorithms and parametric results," *Comp Biol Med*, vol. 36, pp. 448-464, 2006.

[8] V. P. Antipas, G. S. Stamatakos, N. K. Uzunoglu, D. D. Dionysiou, R. G. Dale, "A spatiotemporal simulation model of the response of solid tumors to radiotherapy in vivo: parametric validation concerning oxygen enhancement ratio and cell cycle duration," *Phys Med Biol*, vol. 49(8), pp. 1485 – 1504, 2004.

[9] D. D. Dionysiou, "Computer simulation of in vivo tumor growth and response to radiotherapeutic schemes. Biological optimization of radiation therapy by in silico experimentation," Ph.D. dissertation, Dept. Elect. And Comp. Eng., National Techn Univ., Athens, 2004. (in greek)

[10] S. G. Giatili, G. S. Stamatakos, D. D. Dionysiou, E. A. Kolokotroni, E. C. Georgiadi, "Geometrical and Mechanical Aspects of Tumor Growth and Response to Chemotherapeutic Schemes in the Context of the ACGT Oncosimulator," in Proc. *3rd IARWISO*, Istanbul, Turkey, 2008, pp. 35-37.

[11] E. C. Georgiadi, G. S. Stamatakos, N. M. Graf, E. A. Kolokotroni, D. D. Dionysiou, A. Hoppe, N. K. Uzunoglu, "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," in Proc. *BIBE 2008*, Athens, Greece, 2008. (CFP08266, ISBN: 978-1-4244-2845-8, Library of Congress: 2008907441, Paper No. BE-2.1.2.)

[12] D. D. Dionysiou, G. S. Stamatakos, D. Gintides, N. Uzunoglu, K. Kyriaki, "Critical Parameters Determining Standard Radiotherapy Treatment Outcome for Glioblastoma Multiforme: A Computer Simulation," *The Open Biomedical Engineering Journal*, vol. 2, pp. 43-51, 2008.

The ACGT *Oncosimulator*: from Conceptualization to Development via Multiscale Cancer Modeling

Georgios Stamatakos, Dimitra Dionysiou, Eleni Georgiadi, Eleni Kolokotroni, Stavroula Giatili, Alexander Hoppe, Christine Desmedt, Aran Lunzer, Marius Erdt, Julien Jacques, Juliusz Pukacki, Robert Belleman, Paul Melis, Alberto d'Onofrio, Francesca M. Buffa, Brecht Claerhout, Stefan Rueping, Konstantinos Marias, Manolis Tsiknakis and Norbert Graf

***Abstract* - This short paper provides a brief outline of the main components and the developmental and translational process of the ACGT *Oncosimulator*. The *Oncosimulator* is an integrated software system simulating *in vivo* tumor response to therapeutic modalities within the clinical trial environment. It aims at supporting patient individualized optimization of cancer treatment. The four dimensional simulation module embedded in the *Oncosimulator* is based on the multiscale, *top-down*, discrete entity – discrete event cancer simulation approach and has been specified for the cases of nephroblastoma and breast cancer. Chemotherapeutic treatment in the neoadjuvant setting according to protocols included in the SIOP 2001/GPOH and the TOP clinical trial respectively is considered. The technology modules of the *Oncosimulator* include i.a. anonymized / pseudonymized multiscale data handling (including imaging, histopathological, molecular, clinical and treatment data), image processing and molecular/histopathological data preprocessing, invocation of code execution via an intelligent portal ("RecipeSheet"), execution of the code on either a cluster or grid, collection and visualization of the predictions and numerical analysis (including sensitivity) of the simulation model. Exploratory analyses have revealed the importance of critical model parameters such as the symmetric division probability, the dormancy probability of a newborn tumor cell and the cell kill probability for the response of a solid tumor to chemotherapeutic treatment. Furthermore, plausible parameter values have been adapted to real multiscale clinical trial data in a consistent way, thus supporting the predictive potential of the *Oncosimulator*. Clinical adaptation and validation are in progress. Completion of these time demanding phases is expected to lead to the clinical translation of the system. In parallel, the *Oncosimulator* simulation component is currently expanded in order to allow studying the immune system response to cancer as well as related phenomena and treatment techniques.**

Manuscript received May 5, 2010. (This work has been supported in part by the European Commission under the projects "ACGT: Advancing Clinicogenomic Trials on Cancer" (FP6-2005-IST-026996)).

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

Dimitra Dionysiou (dimdio@esd.ece.ntua.gr), Eleni Kolokotroni (ekolok@mail.ntua.gr), , Eleni Georgiadi (egeorg@ics.forth.gr) and Stavroula Giatili (giatili@otenet.gr) are with the Institute of Communication and Computer Systems, National Technical University of Athens, 9 Iroon Polytechniou, GR 157 80, Zografos, Greece.

Norbert Graf (Norbert.Graf@uniklinikum-saarland.de) and Alexander Hoppe (Alexander.Hoppe@uniklinikum-saarland.de) are with the University Hospital of the Saarland, Paediatric Haematology and Oncology, D-66421 Homburg, Germany.

Christine Desmedt (christine.desmedt@bordet.be) is with the Institut Jules Bordet - Centre des Tumeurs, rue Héger Bordet 1, 1000 Bruxelles, Belgium.

Aran Lunzer (aran@meme.hokudai.ac.jp)is with Hokkaido University, Sapporo, Japan.

Marius Erdt (marius.erdt@igd.fraunhofer.de) is with the Fraunhofer Institute, Darmstadt, Germany

Julien Jacques (Julien.Jacques@irisa.fr) is with the INRIA, Rennes, France.

Juliusz Pukacki (pukacki@man.poznan.pl) is with Poznan Supercomputing and Networking Center (PSNC), Poznan, Poland.

Robert Belleman (R.G.Belleman@uva.nl) and P. Melis (paul@science.uva.nl) are with the University of Amsterdam, The Netherlands.

Alberto d'Onofrio (alberto.donofrio@ieo.it) is with the Istituto Europeo di Oncologia, Milano, Italy.

Francesca M. Buffa (francesca.buffa@imm.ox.ac.uk) is with the Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK.

Brecht Claerhout (brecht@custodix.com) is with Custodix NV, Verlorenbroodstraat 120b14, Merelbeke, Belgium.

Stefan Rueping (Stefan.rueping@iais.fraunhofer.de) is with Fraunhofer IAIS, Schloss Birlinghoven, 53754 St. Augustin, Germany.

Kostas Marias (kmarias@ics.forth.gr) and Manolis Tsiknakis (tsikanki@ics.forth.gr) are with the Foundation for Research and Technology Hellas, Heaklion, Greece.

I. Introduction

THE aim of the present short paper is to briefly outline the modules and the developmental and translational stages of the *Oncosimulator* developed within the framework of the European Commission and Japan co-funded ACGT integrated project (www.eu-acgt.org, http://eu-acgt.org/acgt-for-you/researchers/in-silico-oncology/oncosimulator.html). The *Oncosimulator* is an integrated software system simulating *in vivo* tumor response to therapeutic modalities in the clinical trial context. It aims at supporting patient individualized

optimization of cancer treatment. The four dimensional simulation module embedded in the *Oncosimulator* is based on the multiscale, *top-down*, discrete entity - discrete event cancer simulation approach developed by the *In Silico* Oncology Group, ICCS, National Technical University of Athens [www.in-silico-oncology.iccs.ntua.gr] [1-11]. It has been specified for the cases of nephroblastoma and breast cancer being chemotherapeutically treated in the neoadjuvant setting according to protocols included in the SIOP 2001/GPOH and the TOP clinical trial respectively. The main difference between the two specific models (apart from the different values of the respective model parameters) lies in the different pharmacokinetics and pharmacodynamics submodules for the drugs / drug combinations administered.

Cellular automata, the generic Monte Carlo technique, novel dedicated algorithms and pharmacokinetic differential equations constitute the mathematical basis of the simulation. A discretizing mesh covers the anatomic area of interest. A system of biological cell clusters included within each geometrical cell of the discretizing mesh lies at the heart of the simulation approach. Critical mechanisms such as tumor expansion or shrinkage and the effects of particular drugs on the tumor under consideration have been incorporated into the model. In order to integrate biological mechanisms acting on different biocomplexity scales (levels) the *summarize and jump* strategy [9] has been adopted.

The following processes-modules constitute the core of the integrated scientific and technological construct of the *Oncosimulator*: (i) anonymized / pseudonymized multiscale data collection (including imaging, histopathological, molecular, clinical and treatment data), (ii) image processing and preprocessing of the rest of data, (iii) description of previous or candidate treatment schemes / schedules, (iv) model and simulation code development, (v) invocation of code execution via an intelligent portal (*RecipeSheet*), (vi) execution of the code on either a cluster or the grid, (vii) collection and visualization of the predictions and numerical analysis (including sensitivity analysis) of the simulation model, (viii) evaluation of the predictions and system optimization, (ix) clinical validation of the system within the clinical trial context, (x) optimization of the system, (xi) eventual future translation of the *Oncosimulator* into the clinical practice to serve as a patient individualized treatment optimization support system.

Since the model reflects the natural instability of cancer regarding small changes in a limited number of critical parameters, *virtual cancer patient prototypes* are currently being developed in order to take this behavior into account. Virtual cancer patient prototypes are constructed by combining representative plausible values of critical model parameters by taking into account the molecular, histological and imaging characteristics of a given tumor. Thus each patient prototype is in fact a cluster of combinations of plausible parameter values. By running the simulation code for all candidate treatment schemes for a virtual cancer patient prototype the user can rank the schemes based on the tumor response (e.g. overall cell kill or tumor shrinkage) they produce *in silico*. Thus the treatment scheme/schedule that induces the best tumor response in the majority of parameter value combinations for a given tumor is tentatively assumed to be the optimal one, provided that restrictions imposed by normal tissue side effects allows its clinical application.

It is pointed out that the ACGT *Oncosimulator* constitutes a global novelty.

II. The Medical Problems Addressed

In the case of nephroblastoma, the simulation algorithms address the cases of preoperative chemotherapy for unilateral stage I-III tumors treated in the framework of the clinical trials addressed by ACGT (SIOP 2001/GPOH) with a combination of actinomycin-D and vincristine (Fig. 1).

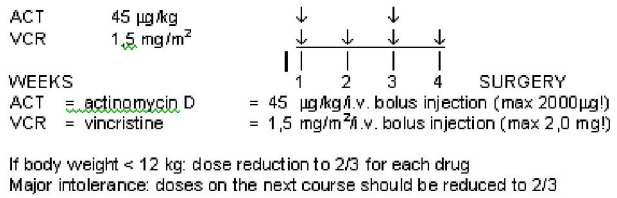


Fig. 1. The simulated treatment protocol for Wilms' tumors.

In the case of breast cancer, the simulation algorithms address the cases of primary chemotherapy ("neo-adjuvant" chemotherapy) with single-agent epirubicin (100 mg/m^2 i.v. once every 3 weeks for 4 consecutive cycles) for early breast cancer patients in the framework of the ACGT clinical trials (TOP trial) (Fig.2).

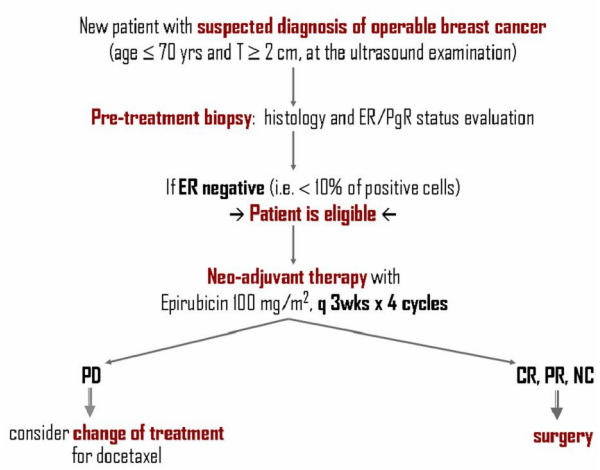


Fig. 2. The simulated treatment protocol for early breast cancer tumors.

III. Indicative Simulation Results

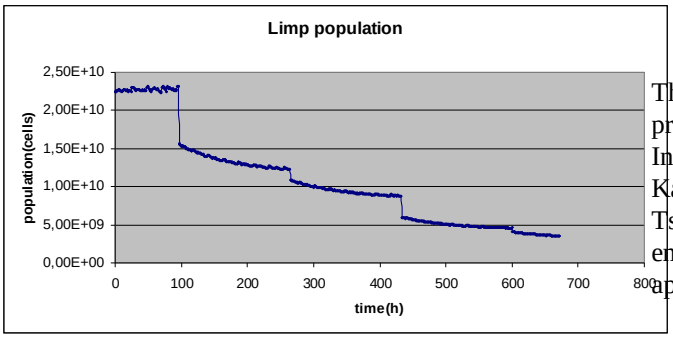


Fig. 3 Simulated time course of the limited mitotic potential (Limp) cell population before, during and after the chemotherapeutic treatment in the case of a typical real nephroblastoma tumor.

Fig.3 shows a typical simulation prediction regarding nephroblastoma (Wilms tumor).

Exploratory analyses have revealed the importance of critical model parameters such as the symmetric division probability (i.e. the probability of a stem cell to give birth to two stem cells instead of a stem and a progenitor cell), the dormancy probability of a newborn tumor cell, the cell cycle duration and the cell kill probability for the response of a solid tumor to chemotherapeutic treatment. Furthermore, plausible parameter values have been adapted to real multiscale clinical trial data (e.g. 20 multiscale data sets in the case of nephroblastoma) in a consistent way thus supporting the predictive potential of the *Oncosimulator*. Clinical adaptation and validation of the ACGT *Oncosimulator* are in progress. An eventually successful completion of this time demanding phase is expected to lead to the clinical translation of the system [12]. In parallel the *Oncosimulator* is currently expanded in order to provide a platform for the study of the immune system response to cancer and related phenomena and/or treatment techniques.

IV. Conclusions

The clinically oriented character of the ACGT *Oncosimulator*, its flexible *top-down*, discrete entity-discrete event simulation philosophy as well as its preliminary clinical adaptation outcome support its potential for future clinical translation. Following the completion of a thorough numerical exploration and a strict clinical validation process, the *Oncosimulator* is expected to serve as treatment optimization system in the patient individualized context. Before reaching that stage utilization of the *Oncosimulator* as a multiscale cancer research tool is an obvious possibility.

Acknowledgments

The authors acknowledge the highly constructive feedback provided by the ACGT external advisors Professor David Ingram, University College London, Professor Olle Björk, Karolinska University, Dr Luca Toldo, and Dr Elena Tsiporkova. They also acknowledge the strong encouragement provided by the European Commission appointed ACGT project officer Dr Ragnar Bergström.

References

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

[2] Dionysiou D.D., 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 , 1-20 , 2004

[3] G.S.Stamatakos, D.D.Dionysiou, N.M.Graf, N.A.Sofra, C.Desmedt, A.Hoppe, N.Uzunoglu, M.Tsiknakis , 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 , Proceedings of the 29th Annual International Conference of the IEEE EMBS Cite Internationale, August 23-26, SuB07.1: 6628-6631 , Lyon, France , 2007

[4] Dionysiou, D.D., G.S. Stamatakos, D. Gintides, N. Uzunoglu K. Kyriaki. 2008. Critical parameters determining standard radiotherapy treatment outcome for glioblastoma multiforme: a computer simulation. Open Biomed Eng J 2: 43-51.

[5] Graf, N., A. Hoppe , E. Georgiadi, R. Belleman, C. Desmedt, D. Dionysiou, M. Erdt , J. Jacques, E. Kolokotroni, A. Lunzer, M. Tsiknakis and G. Stamatakos. 2009. "*In silico* oncology" for clinical decision making in the context of nephroblastoma. Klin Paediatr 221: 141-149.

[6] Kolokotroni E.A.,, G. S. Stamatakos, D. D. Dionysiou, E. Ch. Georgiadi, Ch. Desmedt, N. M. Graf, "Translating Multiscale Cancer

Models into Clinical Trials: Simulating Breast Cancer Tumor Dynamics within the Framework of the "Trial of Principle" Clinical Trial and the ACGT Project.," Proc. 8th IEEE International Conference on Bioinformatics and Bioengineering (BIBE 2008), Athens, Greece, 8-10 Oct. 2008. IEEE Catalog Number: CFP08266, ISBN: 978-1-4244-2845-8, Library of Congress: 2008907441, Paper No. BE-2.1.1, length: 8 pages (in electronic format). 2008

[7] Georgiadi E. Ch., G. S. Stamatakos, N. M. Graf, E. A. Kolokotroni, D. D. Dionysiou, A. Hoppe, N. K. Uzunoglu, "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," Proc. 8th IEEE International Conference on Bioinformatics and Bioengineering (BIBE 2008), Athens, Greece, 8-10 Oct. 2008. IEEE Catalog Number: CFP08266, ISBN: 978-1-4244-2845-8, Library of Congress: 2008907441, Paper No. BE-2.1.2, length: 8 pages (in electronic format). 2008

[8] Giatili S.G., G. S. Stamatakos, D. D. Dionysiou, E. A. Kolokotroni, E. Ch. Georgiadi, Geometrical and Mechanical Aspects of Tumor Growth and Response to Chemotherapeutic Schemes in the Context of the ACGT Oncosimulator," Proceedings of the 3rd International Advanced Research Workshop on In Silico Oncology, Istanbul, Turkey, Sept. 23-24, 2008. Edited by G. Stamatakos and D. Dionysiou pp.35-37, 2008

[9] Stamatakos G. and D. Dionysiou. 2009. Introduction of hypermatrix and operator notation into a discrete mathematics simulation model of malignant tumor response to therapeutic schemes *in vivo*. Some operator properties. Cancer Informatics 2009:7 239–251.

[10] Stamatakos, G.S., E. A. Kolokotroni, D. D Dionysiou, E. C Georgiadi, C. Desmedt, An advanced discrete state - discrete event multiscale simulation model of the response of a solid tumor to chemotherapy. Mimicking a clinical study. J.Theor. Biol. *In press* (2010).

[11] Stamatakos G. 2010. *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. Chapman & Hall/CRC, Boca Raton, Florida, USA, 2010 (*in press*) ISBN: 9781439814406.

[12] N.Graf, *In Silico* Oncology Part II: Clinical Requirements. in T. S. Deisboeck and G. S. Stamatakos Eds, Multiscale Cancer Modeling. Chapman & Hall/CRC, Boca Raton, Florida, USA, 2010 (*in press*) ISBN: 9781439814406.

Breast Cancer Modeling in the Clinical Context: Parametric Studies

Eleni A. Kolokotroni, Dimitra D. Dionysiou, Eleni Ch. Georgiadi, Nikolaos K. Uzunoglu, and Georgios S. Stamatakos

***Abstract*— In the present paper, the dynamic behavior of a clinically-oriented simulation model of breast tumor response to chemotherapy is investigated. The model incorporates various biological processes such as cycling of proliferating cells, quiescence, differentiation and cell death. Indicative results drawn from an extensive parametric analysis of the model are presented.**

I. Introduction

A considerable number of computational models aiming at predicting the response of tumors to various chemotherapeutic schedules have been proposed in literature. For such a model to be eventually translated into clinical practice, a thorough study of its dynamic behavior is a sine qua non prerequisite. A discrete-state, multi-scale simulation model of breast cancer free growth and response to chemotherapy, *in vivo*, has been developed by the *In Silico* Oncology Group (ISOG), National Technical University of Athens, [1],[2]. In this paper, indicative results of an extensive parametric investigation of the model are presented.

II. The Simulation Model

A detailed description of the simulation model can be found in previous publications of the *In Silico* Oncology group [1-4]. The model is based on the consideration of a discrete time and space stochastic cellular automaton, representing the tumor region. More specifically, the tumor region can be considered as a grid of "geometrical cells" (GCs, the elementary volume of the grid). Each GC corresponds to a cluster of heterogeneous cells found in various states. Specific rules regulate the transition between these states, as well as cell movement throughout the tumor volume; the aim is a realistic, conformal to the initial shape of the tumor, simulation of expansion and shrinkage, in the cases of untreated tumor growth and chemotherapy treatment, respectively.

Free growth: The adopted cytokinetic model (Fig.1) includes critical model parameters that have been studied in the present work. It incorporates the biological mechanisms of cell cycling, quiescence, differentiation and loss. Tumor sustenance is attributed to the presence of a minority cell population that exhibits stem cell like properties. Specifically, cancer stem cells have the ability to preserve their own population, as well as give birth to cells of limited mitotic potential (LIMP cells) that follow the path towards terminal differentiation (DIFF cells). A proliferating tumor cell (stem or LIMP) passes through the successive cell cycle phases. Stem, LIMP, DIFF, apoptotic and necrotic cells represent distinct cell categories of the model. Phases within or out of the cell cycle (G1, S, G2, M, G0) constitute different states in which cells may be found. After the completion of mitosis a fraction of newborn cells will enter the dormant phase, whereas the rest will continue to cycle. Transition to quiescence (dormant, G0, phase) and "awakening" of dormant cells are regulated by local metabolic conditions. All cell categories may die through spontaneous apoptosis. However, for dormant and differentiated cells necrosis is the main cell loss mechanism caused by inadequate nutrients' and oxygen supply.

Treatment: When a tumor is chemotherapeutically treated, a fraction of cancerous proliferating cells are lethally hit by the drug. These cells enter a rudimentary cell cycle that leads to apoptotic death through a cell cycle phase depending each time on the specific chemotherapeutic agent. The effect of the drug is considered instantaneous at the time of its administration.

Manuscript received May 5, 2010. This work was supported in part by the European Commission under the projects "ACGT: Advancing Clinicogenomic Trials on Cancer" (FP6-2005-IST-026996), and "ContraCancrum : Clinically Oriented Translational Cancer Multilevel Modeling" (FP7-ICT-2007-2-223979.)

E. A. Kolokotroni (ekolok@iccs.gr), D. Dionysiou (dimdio@esd.ece.ntua.gr), E. Georgiadi (egeorg@ics.forth.gr), N. Uzunoglu (nnap@otenet.gr) and G. Stamatakos (corresponding author, phone: +30-210-772-2287, fax: +30-210-772-3557, gcestam@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.

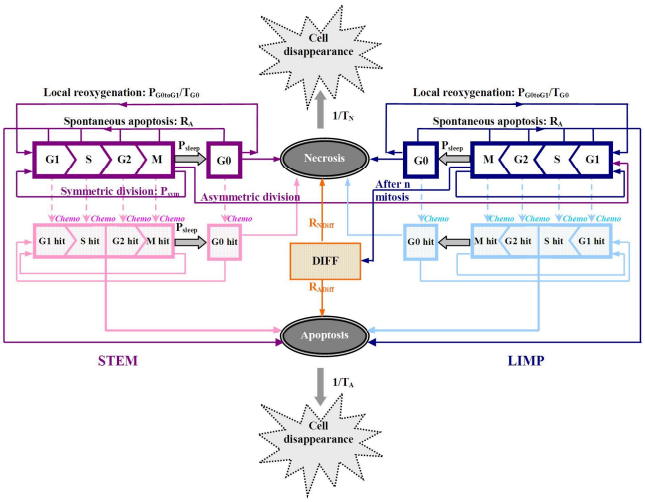


Fig 1. General cytokinetic model for tumor response to chemotherapy. STEM: stem cells. LIMP: Limited proliferative potential cells. DIFF: terminally differentiated cells. G1: Gap 1 phase. S: DNA synthesis phase. G2: Gap 2 phase. M: Mitosis phase. G0: Dormant, resting phase. Chemo: Chemotherapeutic treatment. Hit: Cells lethally hit by the drug.P_{G0toG1}: fraction of dormant cells that reenter the cell cycle, R_A: spontaneous apoptosis rate, P_{sleep}: fraction of newborn cells that enter G0, P_{sym}: fraction of stem cells that perform symmetric division, T_{G0}: duration of dormant phase, T_A: duration of apoptosis, T_N : duration of necrosis, R_{NDiff} : necrosis rate of differentiated cells, R_{ADiff}: apoptosis rate of differentiated cells (see also Table I).

III. Parametric Investigation

The investigation of the effect of each input parameter (Table I) on the behavior of the model is of great importance for a number of reasons, including: detecting biologically irrelevant tumor behavior [1],[2], narrowing the acceptable value range of the model parameters, correlating these parameters with kinetic quantities reported in literature (e.g. doubling time, growth fraction, differentiation grade etc.) and determining the sensitivity or robustness of the output in relation to uncertainties in input data. Indicative results of the corresponding parametric studies are discussed here. The case of a macroscopically homogeneous spherical tumor is considered.

Population initialization: The model aims at simulating the evolution of already fully developed tumors. A reasonable assumption is that such a macroscopically visible tumor can be considered to have reached a state of balanced growth, whereby the population fraction of the various cell categories remains steady over time; this assumption can be made rather safely when considering small time intervals compared to the tumor's lifetime (such as the duration of a chemotherapeutic course), unless a particular perturbation takes place (such as a mutation which transiently changes the kinetic profile of the tumor). An automatic method has been developed that calculates the fraction of each cell population that the cancerous system tends to establish for specific code parameters. These values are fed as an input during the tumor initialization process [1],[2].

Exponential growth: A tumor characterized by a) constant transition rates and phases' durations over space and time and b) population balanced growth, exhibits an exponential free growth pattern. As noted above, such an approximation (constant transitions rates) can be considered applicable for a short time interval.

TABLE I
SUMMARY OF CODE INPUT PARAMETERS AND THEIR ASSIGNED VALUES FOR THE RESULTS PRESENTED IN THE PRESENT WORK UNLESS OTHERWISE STATED

Symbol	Description	Value
*$\mathbf{T_c}$	**Cell cycle duration**	60h†
*$\mathbf{T_{G0}}$	**Duration of dormant phase**	96h
‡T_N	Time needed for necrosis products to disappear from the tumor	120h
‡T_A	Time needed for apoptosis products to be removed from the tumor	6h
N_{LIMP}	Number of mitosis performed by LIMP cells before they become terminally differentiated	7
$\mathbf{R_A}$	**Apoptosis rate of stem/LIMP cells**	0.001
R_{NDiff}	Necrosis rate of differentiated cells	0.004
R_{ADiff}	Apoptosis rate of differentiated cells	0.001
‡$\mathbf{P_{G0toG1}}$	**Fraction of dormant cells that re-enter the cell cycle**	0.01
‡$\mathbf{P_{sleep}}$	**Fraction of cells that enter the G0 phase following mitosis**	0.328
‡$\mathbf{P_{sym}}$	**Fraction of stem cells that perform symmetric division**	0.6
CKR	Cell Kill Ratio	0.5

* Phase durations can be defined separately for the stem and the LIMP tumor cell category.
‡ Defined separately for the proliferating and the necrotic region of the tumor, for spatially inhomogeneous tumor cases.
† Based on literature breast cancer cell cycle duration can vary from 23h to 90h. An intermediate value of 60h is considered.
Bold font denotes parameters related to the stem cell category.

Volume doubling time (T_d): A thorough study of the effect of each model input parameter on T_d has been performed (results not shown). The results reveal that the volume doubling time is affected only by the parameters related to the cell category that is responsible for tumor sustenance, i.e the stem cell category (parameters denoted in bold in Table 1). Depending on the values attributed to these parameters it is possible to simulate tumors of different aggressiveness. Furthermore there are certain 'forbidden' value combinations of these parameters that lead to self-diminishing over time tumors. Based on a compartmental analysis of our model, the following equation for the growth rate, α, of the tumor is obtained:

$$e^{(\alpha+R_A)T_C} = \left(1+P_{sym}\right)\left(1-P_{sleep}+P_{sleep}\frac{P_{G0toG1}/T_{G0}}{R_A+1/T_{G0}+\alpha}\right) \quad (1)$$

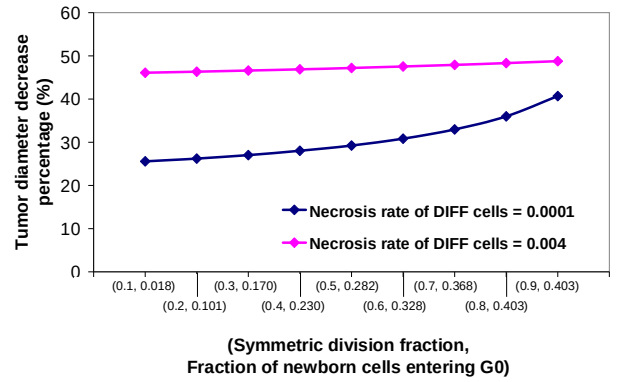
For monotonically growing tumors α must be positive.

Chemotherapeutic response: As anticipated more

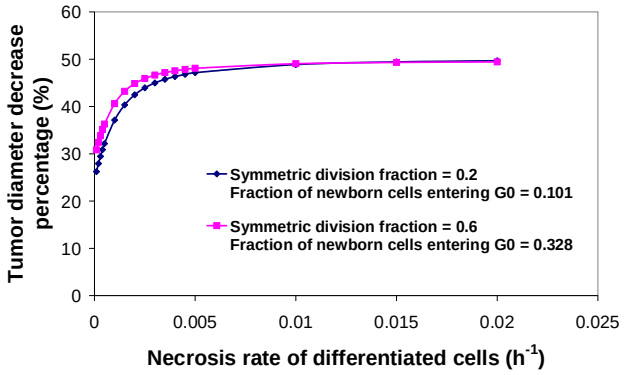
aggressive tumors, i.e. tumors characterized by short volume doubling times, exhibit a worse chemotherapeutic response (results not shown). Different combinations of model parameters may result in the same doubling time. It is interesting to examine the simulation outcome when a given chemotherapeutic schedule is applied on these different implementations of tumors with the same doubling time.

In Fig. 2 we investigate the joint effect of three model parameters - symmetric division fraction, P_{sym}, fraction of newborn cells that enter the G0 phase, P_{sleep} and necrosis rate of differentiated cells, R_{NDiff}, - on tumor shrinkage as a result of chemotherapeutic treatment. In the upper panel the parameters P_{sym} and P_{sleep} vary in combination so as to maintain a constant T_d (equal to 100 days). On the contrary we can vary R_{NDiff} without affecting T_d. We observe that (a) tumors characterized by higher values of the pair P_{sym} - P_{sleep} or higher R_{NDiff} have better response to therapy, (b) the model's output is robust to variation of P_{sym} - P_{sleep} pair when differentiated cells have a short lifespan (high values of R_{NDiff}) and (c) for high values of R_{NDiff} tumor shrinkage is not affected by this parameter.

As the chemotherapeutic agent considered in these studies (epirubicin) targets proliferating and G0 cells (stem and LIMP), there is a delay in the appearance of the effect of therapy on the rest of the cell categories (differentiated and dead cells). This delay is small (and hence the therapeutic response better), when the lifespan of the differentiated and dead cell categories (the reciprocal of the corresponding removal rates) is short and/or their portion in the bulk of the tumor is low. Higher values of R_{NDiff} are associated with shortly lived differentiated cells, whereas higher values of the parameters P_{sym} and P_{sleep} result in higher fractions of proliferating and G0 cells and hence lower fractions of differentiated and dead cells. The effect of the variation of the above parameters on therapeutic outcome is insignificant when the delay is small enough compared to the period between two drug administrations. These results are in accordance with clinical experience according to which tumors with higher growth fractions exhibit better response to chemotherapy.



(a)



(b)

Fig. 2. The effect of (a) symmetric division fraction (P_{sym}) - fraction of newborn cells that enter G0 phase (P_{sleep}) and (b) necrosis rate of differentiated cells (R_{NDiff}) on tumor diameter. The values assigned to the model parameters are included in Table I. The volume doubling time of the tumor is equal to 100 days in all cases. A homogeneous spherical tumor of an initial diameter equal to 10mm and a typical density of 10^6 biological cells/mm is considered. Epirubicin is assumed to be administrated orally (bolus administration). The dose fragmentation scheme simulated is once every 3 weeks for 4 consecutive cycles. The final tumor diameter is measured 1 week after the last drug administration.

IV. Conclusions

An advanced simulation model of breast cancer response to chemotherapy has been developed. Prior to its clinical use the model must undergo a strict validation and clinical optimization procedure. A critical first step towards the model's validation is to elucidate the effect and interplay of critical parameters on model behavior. An extensive parametric study in order to better understand the model's sensitivity to input code parameters variation has been performed. Indicative results have been presented in the present paper. Quantitative validation and optimization based on pertinent clinical and laboratory data (e.g. tumor size reduction after therapy, gene expression profile etc.) is under way in the framework of the TOP trial (http://clinicaltrials.gov/ct2/show/NCT00162812.)

REFERENCES

[1] E. A. Kolokotroni, G. S. Stamatakos, D. D. Dionysiou, E. Ch. Georgiadi, Ch. Desmedt, N. M. Graf, "Translating Multiscale Cancer Models into Clinical Trials: Simulating Breast Cancer Tumor Dynamics within the Framework of the "Trial of Principle" Clinical Trial and the ACGT Project.," in proc. *8th IEEE International Conference on Bioinformatics and Bioengineering (BIBE 2008)*, Athens, Greece, 8-10 Oct. 2008. IEEE Catalog Number: CFP08266, ISBN: 978-1-4244-2845-8, Library of Congress: 2008907441, Paper No. BE-2.1.1, length: 8 pages (in electronic format). 2008

[2] G. S. Stamatakos, E. A. Kolokotroni,, D. D. Dionysiou, E. Ch. Georgiadi, C. Desmedt, "An advanced discrete state – discrete event multiscale simulation model of the response of a solid tumor to chemotherapy. Mimicking a clinical study," *Journal of Theoretical Biology*, accepted, to be published.

[3] D. D. Dionysiou, G. S. Stamatakos, N. K. Uzunoglu, N.K., K. S. Nikita, "A computer simulation of in vivo tumor growth and response to radiotherapy: New algorithms and parametric results," *Comp. Biol. Med.* vol. 36, pp. 448–464, 2006.

[4] S. G. Giatili, G. S. Stamatakos, D. D. Dionysiou, E. A. Kolokotroni, E. Ch. Georgiadi, "Geometrical and Mechanical Aspects of Tumor Growth and Response to Chemotherapeutic Schemes in the Context of the ACGT Oncosimulator," in proc. *3rd International Advanced Research Workshop on In Silico Oncology*, Istanbul, Turkey, Sept. 23-24, 2008. Edited by G. Stamatakos and D. Dionysiou pp.35-37

Discrete Event Based Modeling of Nephroblastoma. Sensitivity Considerations

Eleni Ch. Georgiadi, Dimitra D. Dionysiou, Eleni Kolokotroni, Nikolaos K. Uzunoglu, Norbert Graf, and Georgios S. Stamatakos

***Abstract*—In this paper indicative sensitivity studies performed with an already developed simulation model of nephroblastoma response to chemotherapy are presented. The sensitivity analyses reported involve model parameters with a critical effect on the following aspects of the model: tumor initialization, tumor growth and tumor response to chemotherapy.**

I. Introduction

AN advanced discrete multilevel simulation model of nephroblastoma growth and response to chemotherapy has been recently developed by our group [1] and is now in the process of clinical validation within the context of the SIOP 2001 / GPOH clinical trial [2]. The trial addresses the case of preoperative chemotherapy for unilateral stage I-III tumors treated with a combination of actinomycin-D and vincristine. Initial observations implying qualitative agreement of the model's behavior with the corresponding clinical trial experience have been already presented [1],[2]. A thorough sensitivity and parameter interdependence analysis aiming at fully determining the model's behavior throughout the value space of its parameters has been carried out; representative results are presented in the present paper.

II. Identifying Critical Model Parameters

A reference value is defined for each model parameter based on pertinent literature and the dictates of accumulated basic science and clinical experience (Table I). In order to identify critical parameters in terms of their effect on the simulation outcome, three exploratory runs are performed for each parameter under study: one with the reference value and the other two by varying this value by 5% and -5%, while keeping the values of the remaining model parameters constant. The sorting criterion is the magnitude of deviation (absolute maximum deviation) from the final tumor volume obtained when using the reference value of each parameter (Fig.1).

As shown in Fig. 1, the parameters having the maximum effect on the tumor's evolution and the outcome of chemotherapy are:

a. Fraction of cells in the proliferative/necrotic region that will enter G0 following mitosis
b. Fraction of stem cells in the proliferative/necrotic region that divide symmetrically.
c. Fraction of any non differentiated cell subpopulation that undergoes spontaneous apoptosis per hour (h^{-1})
d. Fraction of any dormant stem or limp cell subpopulation that re-enters cell cycle per hour (h^{-1})
e. Cell cycle duration

TABLE I : The parameters modeling the growth rate of a tumor: definition and reference value. The parameters T_A, P_{sym}, P_{sleep} can be assigned different values in the proliferative and necrotic regions of the tumor.

Symbol	*Definition*	*Reference Value*
T_c	Cell cycle duration *(in hours)*	23.0
T_{G0}	Maximum G0 phase duration before a stem/limp cell dies through necrosis *(in hours)*	96
T_N	Time (*in hours*) before necrosis products are eliminated	20
T_A	Time (*in hours*) before necrosis products are eliminated	6
R_A	Fraction of any non differentiated cell subpopulation that undergoes spontaneous apoptosis per hour (h^{-1})	0.001
R_{ADiff}	Fraction of any differentiated cell subpopulation that undergoes spontaneous apoptosis per hour (h^{-1})	0.003
R_{NDiff}	Fraction of any differentiated cell subpopulation that enter necrosis per hour (h^{-1})	0.001
R_{G0toG1}	Fraction of any dormant stem or limp cell subpopulation that re-enters cell cycle per hour (h^{-1})	0.01
N_{LIMP}	Number of mitoses that a limp (progenitor) cell undergoes before becoming terminally differentiated	3
P_{sym}	Fraction of stem cells that divide symmetrically.	0.45
P_{sleep}	Fraction of cells that enter G0 following mitosis	0.28

Manuscript received May 15, 2010. This work was supported in part by the European Commission under the projects "ACGT: Advancing Clinicogenomic Trials on Cancer" (FP6-2005-IST-026996) and "ContraCancrum: Clinically Oriented Translational Cancer Multilevel Modeling" (FP7-ICT-2007-2-223979.)

E. C. Georgiadi (egeorg@ics.forth.gr) , D. D. Dionysiou (dimdio@esd.ece.ntua.gr), E. Kolokotroni (ekolok@mail.ntua.gr), N. K. Uzunoglu (nuzu@cc.ece.ntua.gr) , G. S. Stamatakos(corresponding author gestam@central.ntua.gr, phone: +30-210-772 22 87; fax: +30-210-772 3557) are with the *In Silico* Oncology Group, Institute of Communication and Computer Systems, National Technical University of Athens, Iroon Polytechniou 9, Zografos, GR-157-80, Athens, Greece.

N. Graf (graf@uks.eu) is with the University Hospital of the Saarland, Paediatric Haematology and Oncology, D-66421 Homburg, Germany 66421.

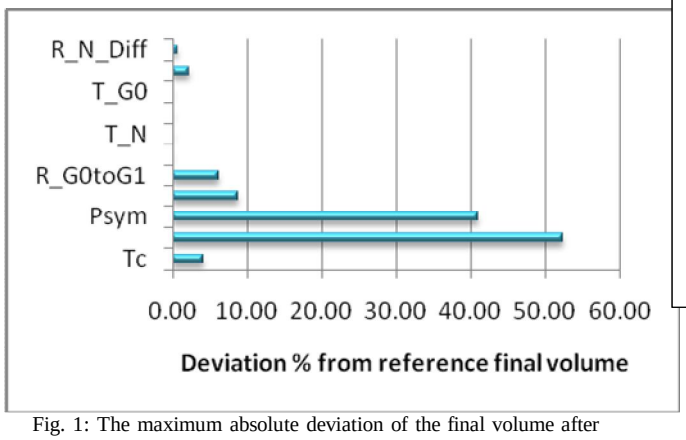


Fig. 1: The maximum absolute deviation of the final volume after therapy from the reference final volume, as calculated by varying the reference value of each parameter by ±5% .

III. SENSITIVITY ANALYSIS CONSIDERATIONS

1) Effect of the model parameters on tumor initialization

As pointed out in previous publications [1],[3], the initial distribution of cells to the various sub-categories has a great impact on the evolution of a simulated tumor. Initialization is performed by placing a small number of stem cells within the mesh. This initial stem cell population evolves over time to produce the remaining cell categories and a state of equilibrium in terms of relative population percentages (expressed as fractions of the total tumor cell population) is ultimately reached for specific cell category and cell phase transition rates. The challenge is to successfully detect the point of equilibrium, and then use the stabilized fractions of the populations for the correct initialization of the tumor [4]. In Fig. 2 the currently used method for tumor initialization is outlined.

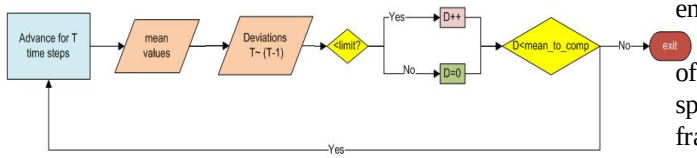


Fig. 2. Flowchart of the tumor's subpopulations initialization method. Advance: Time evolution of the tumor under free growth conditions. T: number of time steps for which mean values are calculated. Mean_to_comp: number of mean values to compare. D: counter. Limit: set to 10^{-7}. Mean value: refers to cell category populations expressed as fractions of the total cell population.

$$mean\_cell\_category\_fraction_i = \sum_{iT}^{(i+1)T-1} \frac{cell\_category\_fraction}{T}$$. Deviations:

$$deviation = \left| \frac{mean\_cell\_category\_fraction_{i+1} - mean\_cell\_category\_fraction_i}{mean\_cell\_category\_fraction_i} \right|$$

The effect of the model parameters on the initialization of the tumor has been studied and indicative results for some of the most important parameters are presented below (Fig. 3, Fig. 4).

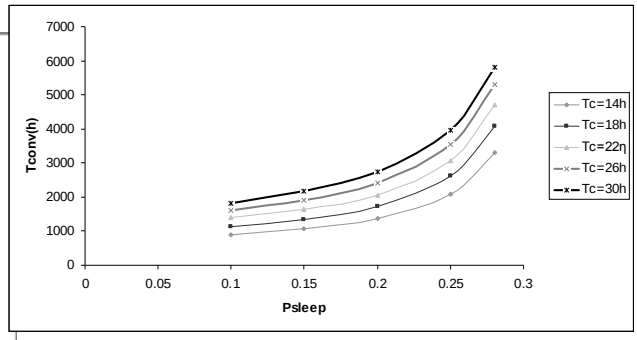


Fig.3: The required convergence time (T_{conv}) for an appropriate initialization of the simulated tumors in relation to the fraction of cells that enter G0 following mitosis (P_{sleep}) for various cell cycle durations (T_c).

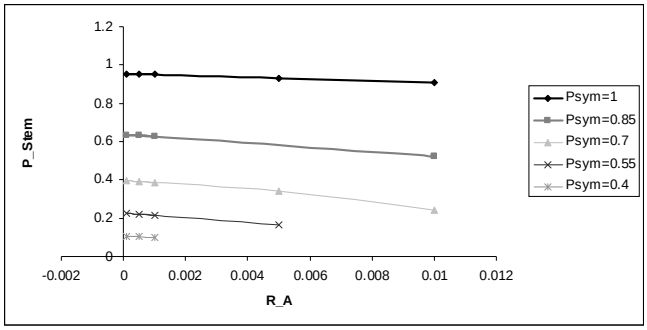


Fig.4: Initial fraction of stem cells (P_Stem) in relation to the spontaneous apoptosis rate (R_A) for various symmetric division fractions (P_{sym}).

As far as the required time for convergence is concerned, tumors with high growth rate reach equilibrium faster. As depicted in Fig. 3, a tumor with a short cell cycle duration needs less time to reach equilibrium; the same holds true for a tumor with a small fraction of cells entering G0 following mitosis.

In Fig. 4 the initial population of stem cells (as fraction of the total cell population) is presented in relation to the spontaneous apoptosis rate for various symmetric division fractions. It is obvious that the symmetric division fraction has a greater effect on the initial stem cell population, whereas the influence of the spontaneous apoptosis rate is limited especially for high symmetric division percentages. The higher the symmetric division percentage, the higher is the initial stem cell fraction, thereby leading to a faster growing tumor.

2) Effect of the model parameters on tumor growth

The effect of the various model parameters on the doubling time of a growing tumor has been studied. As far as the tumor's growth rate is concerned, it should be noted that the simulated time intervals are small enough to consider exponential fitting of the growth curve (1); this exponential growth represents only a segment of the actual Gompertzian growth curve. Careful examination of the pattern of temporal variation of the various cell category populations for limited time intervals reveals that their variation appears to be linear on semi-logarithmic plots, i.e

an exponential behavior seems to be justified. Therefore, the population of a cell category evolves over time according to the equation:

$$N(t) = N_o e^{at} \quad (1)$$

And the doubling time may be retrieved by:

$$T_d = \frac{\ln 2}{\kappa} \quad (2)$$

As an indicative result the doubling time in relation to the fraction of cells that enter G0 following mitosis ("sleep fraction") for various cell cycle durations is presented in Fig. 5. Tumors with longer cell cycle duration have higher doubling time than tumors with short cell cycle duration. This difference in doubling time becomes more noticeable for tumors with higher sleep fractions.

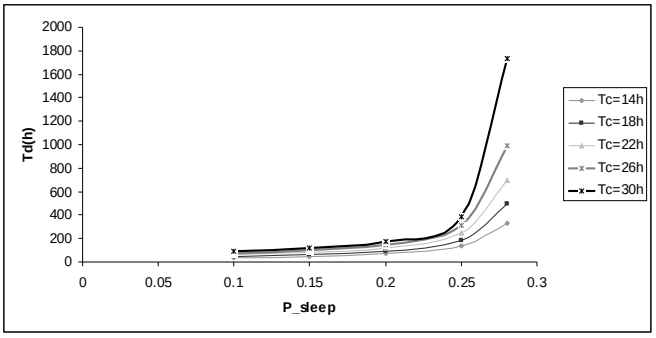


Fig.5: The tumor doubling time (Td) in relation to the fraction of cells that enter G0 following mitosis (P_sleep), for various cell cycle durations.

3) Effect of the model parameters on chemotherapy outcome

The final outcome of chemotherapy treatment depends on the model parameters that define the growth rate of the tumor, as well as the parameters that define the effect of therapy on the tumor. In order to quantify the above effects, the absolute tumor shrinkage (3) has been used:

$$DV = \left| \frac{V_{final} - V_{initial}}{V_{initial}} \right| * 100\% \quad (3)$$

where $V_{initial}$ and V_{final} are the initial and the final tumor volume, respectively.

The effect of the cell kill ratios of actinomycin and vincristine administered according to the SIOP treatment protocol [3],[7] is presented in Fig. 6 and Fig. 7. As expected, higher cell kill ratios lead to greater tumor volume reduction. Vincristine's cell kill ratio has a greater effect on the outcome of chemotherapy, as the treatment protocol specifies 4 sessions of vincristine administration and 2 sessions of actinomycin administration.

IV. CONCLUSION

Selected sensitivity studies performed with an already developed simulation model of Wilms' tumor growth and response to chemotherapy have been presented. Specifically, the parameters of the model that have a great effect on tumor initialization, tumor growth and tumor response to therapy have been detected. Indicative analysis on the effect of some of the basic parameters on the sensitivity quantities is depicted.

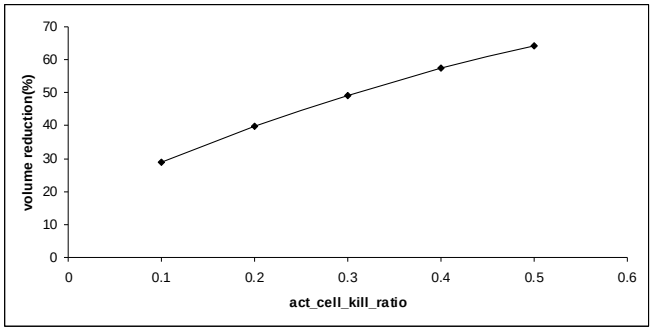


Fig. 6: The reduction of tumor volume in relation to the cell kill ratio of actinomycin.

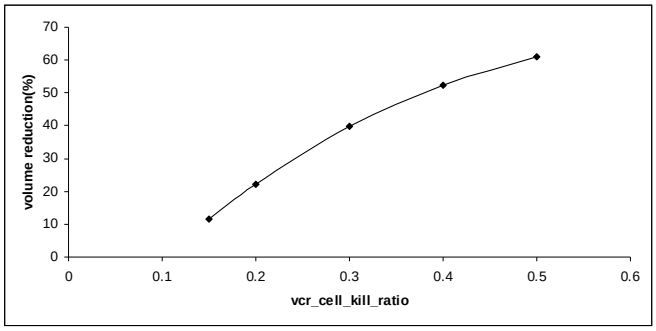


Fig. 7: The reduction of tumor volume in relation to the cell kill ratio of vincristine

REFERENCES

[1] E. Ch. Georgiadi , G. S. Stamatakos, N. M. Graf, E. A. Kolokotroni, D. D. Dionysiou, A. Hoppe, N. K. Uzunoglu, "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," in proc. *8th IEEE International Conference on Bioinformatics and Bioengineering (BIBE 2008*). IEEE Catalog Number: CFP08266, ISBN: 978-1-4244-2845-8, Library of Congress: 2008907441, Paper No. BE-2.1.2, 2008.

[2] N. Graf, A. Hoppe , E. Georgiadi, R. Belleman, C. Desmedt, D. Dionysiou, M. Erdt , J. Jacques, E. Kolokotroni, A. Lunzer, M. Tsiknakis and G. Stamatakos, "*In silico* oncology for clinical decision making in the context of nephroblastoma.," *Klin Paediatr* vol. 221, pp.141-149, 2009.

[3] E. A. Kolokotroni, G. S. Stamatakos, D. D. Dionysiou, E. Ch. Georgiadi, Ch. Desmedt, N. M. Graf, "Translating Multiscale Cancer Models into Clinical Trials: Simulating Breast Cancer Tumor Dynamics within the Framework of the "Trial of Principle" Clinical Trial and the ACGT Project.," in proc. *8th IEEE International Conference on Bioinformatics and Bioengineering (BIBE 2008)*. IEEE Catalog Number: CFP08266, ISBN: 978-1-4244-2845-8, Library of Congress: 2008907441, Paper No. BE-2.1.1, 2008

[4] G. S. Stamatakos, E. A. Kolokotroni,, D. D. Dionysiou, E. Ch. Georgiadi, C. Desmedt, "An advanced discrete state – discrete event multiscale simulation model of the response of a solid tumor to chemotherapy. Mimicking a clinical study," Journal of Theoretical Biology *in press*

Clinically Oriented Translational Cancer Multilevel Modeling

Kostas Marias, Vangelis Sakkalis, Alexandros Roniotis, Ioannis Karatzanis, Georgios Stamatakos, Dimitra Dionysiou, Stavroula Giatili, Nikolaos K. Uzunoglou, Norbert Graf, Rainer Bohle, Eckart Messe, Holger Stenzhorn, Yoo-Jin Kim, Peter V. Coveney, Stefan J. Zasada, Shunzhou Wan, Amos Folarin, Philippe Büchler, Tibault Bardyn, Stefan Bauer, Mauricio Reyes, Gordon J Clapworthy, Enjie Liu, Tomas Bily, Vojtech Bednar, Michal Karasek, Astrid Franz, Rüdiger Grewer, and Jörg Sabczynski

***Abstract*— ContraCancrum means 'Against Cancer' in Latin. This is because the ContraCancrum project aims to pave the way for translating clinically validated multilevel cancer models into clinical practice. The models will assist the clinician to define the optimal therapy for the individual patient taking into consideration all the available clinical data from different scales (molecular to tissue level), modalities (e.g. multimodal cancerimaging) and examinations (e.g. before and after therapy).**

This work was supported by the European Commission under the project "ContraCancrum: Clinical Oriented Translational Cancer Multilevel Modeling" (FP7-ICT-2007-2- 223979).

K. Marias, V. Sakkalis, A. Roniotis, and I. Karatzanis are with the Institute of Computer Science at FORTH, Vassilika Vouton, GR-70013 Heraklion, Crete, Greece (email: kmarias@ics.forth.gr , sakkalis@ics.forth.gr , roniotis@ics.forth.gr , karatzan@ics.forth.gr).

G. Stamatakos, D. Dionysiou, S. Giatili, N. K. Uzunoglou are with the Institute of Communication and Computer Systems, Dept of Electrical and Computer Engineering, Technical University of Athens, Athens, 15780, Greece (email: gestam@central.ntua.gr, dimdio@esd.ece.ntua.gr, giatili@otenet.gr)

N. Graf, H. Stenzhorn is with the University of the Saarland, Department of Pediatric Oncology, 66421 Homburg, Germany (email: Norbert.Graf@uniklinikum-saarland.de, holger.stenzhorn@ifomis.uni-saarland.de)

R. Bohle, Y.-J- Kim is with the University of the Saarland, Department of Pathology, 66421 Homburg, Germany (email: rainer.bohle@uniklinikum-saarland.de , yoo-jin.kim@uniklinikum-saarland.de)

E. Messe, Y.-J- Kim is with the University of the Saarland, Department of Human Genetics, 66421 Homburg, Germany (email: hgemee@uniklinik-saarland.de)

P. V. Coveney, S. J. Zasada, S. Wan are with the Chemistry Department, University College London, London, UK. (e-mails: p.v.coveney@ucl.ac.uk, stefan.zasada@ucl.ac.uk, shunzhou.wan@ucl.ac.uk).

A. Folarin is with the Cancer Institute, University College London, London, UK. (e-mail: a.folarin@ucl.ac.uk).

P. Büchler, T. Bardyn, S. Bauer, M. Reyes are with the Institute for Surgical Technology and Biomechanics, University of Bern, Switzerland, Staufacherstrasse 78, 3014, Switzerland (email: Philippe.Buechler@artorg.unibe.ch, thibaut.bardyn@artorg.unibe.ch, stefan.bauer@istb.unibe.ch, mauricio.reyes@istb.unibe.ch)

G J Clapworthy and E Liu are with the University of Bedfordshire, Luton, LU1 3JU, UK. (e-mail: gordon.clapworthy@beds.ac.uk and enjie.liu@beds.ac.uk).

T. Bily, V. Bednar, M.Karasek are with Charles University in Prague Faculty Mathematics and Physics, Department of Applied Mathematics, Prague 1, 118 00, Czech Republic (e-mail: tomby@kam.mff.cuni.cz, bednar@kam.mff.cuni.cz, michal.karasek@gmail.com).

A. Franz, R. Grewer, and J. Sabczynski are with Philips Technologie GmbH Forschungslaboratorien, Hamburg, Germany (email: Astrid.Franz@philips.com , Ruediger.Grewer@philips.com , Joerg.Sabczynski@philips.com).

I. Introduction

CONTRACANCRUM [1-2] is developing a composite multilevel platform for simulating malignant tumor development and tumor and normal tissue response to therapeutic modalities and treatment schedules. The clinical validation of the models will be achieved by dedicated clinicogenomic studies in gliomas and lung cancer. The results of the simulations will be validated by the real outcome of treatment response using imaging studies. Tumor tissues and blood samples of about 100 patients are being obtained from patients with lung cancer and 50 with glioblastoma, with their informed consent. Tissue samples are used for gene expression profiling. From each patient imaging studies are stored as DICOM files at the time of diagnosis, after surgery and at the end of treatment. The collected data include treatment data, gene array expression data and imaging and are anonymized and uploaded in the ContraCancrum environment for the ContraCancrum simulations. This environment, provides a data warehousing system in which anonymised patient data can be stored. This data falls in to three broad categories: imaging data, stored in the DICOM format, structured clinical records and file based data, generated by the different types of simulation used in the project. The data environment ties together this different types of data through the use of unique patient identifiers. All components of the technical environment including workflow editor and webservices, are accessed through a web portal, which manages access policies and presents data and results to users.

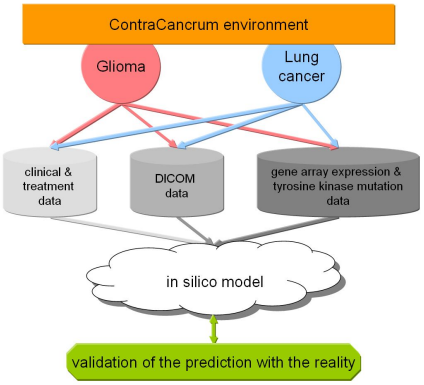


Fig. 1. Schema of both clinical scenarios in ContraCancrum.

II. Project Components

A. Microscopic tumor growth module

Microscopic tumor growth module includes simulations of microscopic mechanisms of tumor growth and response to treatment including i.a. avascular tumor growth, response to treatment, angiogenesis, invasion, and metastasis. This makes it possible to design experiments which are heavy feasible in vivo and in vitro and to enhance our understanding of the tumor dynamics related phenomena on the microscopic level so that refinement of the imageable tumor models can be achieved.

B. Cellular and higher level tumor dynamics module

This module includes the development of a set of multilevel simulation models of tumor growth as well as tumor response to radiotherapy and chemotherapy for the cases of glioblastoma multiforme and lung cancer in the patient individualized context. Both discrete and continuous simulation models of tumor growth out of a single tumor cell as well as tumor response to radiotherapy and chemotherapy are exploited in ContraCancrum.

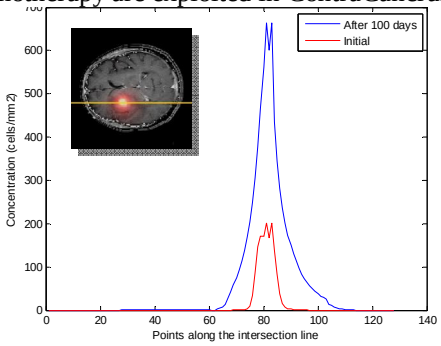


Fig 2. Simulation of free glioblastoma growth

C. The biomechanical simulator

The objective of the biomechanical simulator is to consider the mechanical environment inside and outside the tumor in the model. Information about tumor growth is obtained from the molecular and cellular levels de-veloped in the project and fed into the model. A fully automatic meshing algorithm for the different tissues has been built for this purpose. Based on this mesh, a continuum finite element model is proposed to simulate the tumor, its growth and the mechanical perturbations induced on the surrounding healthy tissues. The mechanical information obtained is then transmitted back to the cellular simulator.

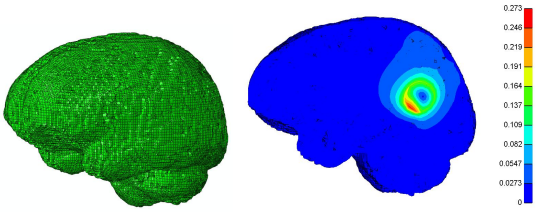


Fig. 3: Biomechanical simulations in brain cancer data

D. Biochemical simulations and molecular determinants of response to therapy

The biochemical simulator employs large scale molecular dynamics (MD) techniques to study the interactions between inhibitors and target proteins in atomistic detail, together with free energy calculations to rank drug binding affinities. Using the biochemical simulator, we have investigated the binding affinities of two tyrosine kinase inhibitors – AEE788 and Gefitinib – to epidermal growth factor receptor (EGFR), an attractive target for anti-lung cancer therapies. The interactions between these inhibitors and the wild-type and mutant EGFR kinase domains are analyzed using multiple short (ensemble) simulations. The simulations reveal how interactions change as a result of the mutations, and account for the molecular basis of drug efficacy. The free energy calculations show that the method is able to rank binding affinities of one drug to multiple EGFR mutants, as well as the efficacy of drugs with respect to a single EGFR sequence.

Another molecular level module aims to provide a statistical a priori model of individual response to therapy. Using drug and radiation sensitivity data, coupled with microarray expression data, we have identified signatures capable of predicting in vitro response to therapy as a measure of cell-survival. A statistical model is constructed based on the correlation between gene expression and therapy induced inhibition of cell growth (GI50). This statistical model is then used to assess a given patient's tumor profile, thus providing an estimation of cellular therapy response. This component additionally defines the means to incorporate molecular information within the context of in silico simulation of patient-specific therapy as part of the ContraCancrum multi-level platform by modifying the cell survival probability within the tissue level component.

E. Image analysis modules

In the case of gliomas for example, different modalities of imaging techniques (e.g. T1 MRI, T2 MRI, Flair MRI, CT, PET etc) are required. Even so, it is very difficult for a radiologist to define the exact margins of bulk tumor. A real example indicating the major problem of defining the

exact margins of a glioma by clinicians is given in Fig. 4. Even with 4 different modalities (T1, T1 enhanced, T2, T2 flair), glioma boundaries are not easily discernible and within each boundary different pathophysiological information is contained. The clinician and the modeller need to get these multi-modal boundaries in order to gather information concerning the individual tumor properties. The image analysis component is providing a variety of powerful tools that enable the clinician to register, segment and extract cancer pathophysiological parameters from multimodal cancer data. Non-rigid registration of images from different modalities has been developed through a method based on Markov Random Fields, while novel segmentation techniques, such as the spatially adaptive active contours, have been integrated in the platform.

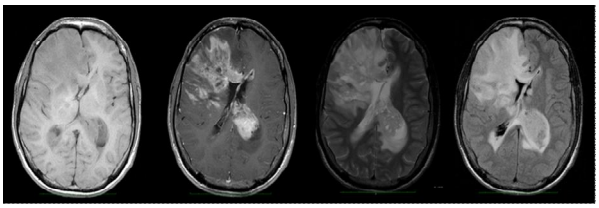


Fig. 4: Different modalities taken with different MRI techniques. From left to right: T1, T1-enhanced, T2, T2 flair

F. ContraCancrum technical environment

The ContraCancrum technical environment has three main components that are accessed through a web portal, which manages access policies and presents data and results to users. The three distinct components are:

- The ContraCancrum Data Environment, which provides a data warehousing system in which anonymised patient data can be stored. This data falls in to three broad categories: imaging data stored in the DICOM format, structured clinical records, and file-based data generated by the different types of simulation used in the project. The data environment ties together these different types of data by the use of unique patient identifiers (and, in the future, through ontologies). This means that a user of the central data repository can query and view multiple data types held either on a single patient or on a population.
- A set of web services, which provide a standard way to access different simulation tools, including the ability to segment medical images and launch simulations on high performance computing resources.
- A workflow engine, which couples together data resources and simulation services, to automate the processing of the different types of data available, and to tie different simulation scales together.

The ContraCancrum user will be able to utilise the project components described above for multi-scale cancer predictive oncology simulations, including microscopic tumor growth simulation, cellular and higher-level tumor dynamics simulation, biomechanical simulations, biochemical simulations and molecular determinants of response to therapy and image analysis modules.

III. CONCLUSION

The major expected potential impact is a substantial contribution to the personalized optimization of cancer treatment strategies. Furthermore, strengthening European biomedical industry through the development of new decision support systems is expected to be achieved. Obviously, the central socio-economic and societal impact is expected to be the alleviation of the multidimensional societal burden of the cancer disease.

REFERENCES

[1] ContraCancrum website: http://www.contracancrum.eu/

[2] Marias, V. Sakkalis, A. Roniotis, C. Farmaki, G. Stamatakos, D. Dionysiou, S.Giatili, N. Uzunoglou, N. Graf, R. Bohle, E. Messe, P. V. Coveney, S. Manos, S. Wan, A. Folarin, S. Nagl, P. Büchler, T. Bardyn, M. Reyes, G. Clapworthy, N. Mcfarlane, E. Liu, T. Bily, M. Balek, M. Karasek, V. Bednar, J. Sabczynski, R. Opfer, S. Renisch, and I. C. Carlsen, "Clinically Oriented Translational Cancer Multilevel Modeling: The ContraCancrum project", World Congress 2009 on Medical Physics and Biomedical Engineering, Sept 7-12, Munich, Germany.

The ContraCancrum *Oncosimulator*: Integrating Biomechanisms Across Scales in the Clinical Context

Georgios S. Stamatakos, Dimitra Dionysiou, Stavroula Giatili, Eleni Kolokotroni, Eleni Georgiadi, Alexandros Roniotis,Vangelis Sakkalis, Peter Coveney, Shunzhu Wan, Steven Manos, Stefan Zasada, Amos Folarin, Philippe Büchler, Tibault Bardyn, Stefan Bauer, Mauricio Reyes, Tomas Bily, Vojtech Bednar, Michal Karasek, Norbert Graf, Rainer M. Bohle, Eckart Meese, Yoo-Jin Kim, Holger Stenzhorn, Gordon Clapworthy, Enjie Liu, Jörg Sabczynski, and Kostas Marias

***Abstract*— This short communication briefly outlines the major components and the integration steps of the *Oncosimulator* that is being developed within the framework of the European Commission funded ContraCancrum project. The *Oncosimulator* is a technologically advanced multiscale tumor growth and treatment response system aiming at supporting patient individualized treatment decisions. An indicative example of the adopted mathematical approaches as well as a simple example of numerical code validation are provided. The document concludes with a short discussion on the characteristics of the major modeling approaches that refer to the cellular and higher biocomplexity levels since the latter constitute the basis for the entire *Oncosimulator* integration.**

Manuscript received May 5, 2010. (This work has been supported in part by the European Commission under the project ContraCancrum: Clinically Oriented Translational Cancer Multilevel Modeling" (FP7-ICT-2007-2-223979)).

G. S. Stamatakos (corresponding author) is with the Institute of Communication and Computer Systems, National Technical University of Athens, 9 Iroon Polytechniou, GR 157 80, Zografos, Greece (phone: +30 210772 2287; fax: +30 2107733557; e-mail: gestam@central.ntua.gr).

D. Dionysiou (dimdio@esd.ece.ntua.gr), S. Giatili (giatili@otenet.gr) E.Kolokotroni (ekolok@mail.ntua.gr) and E. Georgiadi (egeorg@ics.forth.gr) are with the Institute of Communication and Computer Systems, National Technical University of Athens, 9 Iroon Polytechniou, GR 157 80, Zografos, Greece.

K.Marias (kmarias@ics.forth.gr),V. Sakkalis (sakkalis@ics.forth.gr) and A. Roniotis (roniotis@ics.forth.gr) are with the Institute of Computer Science at FORTH, Vassilika Vouton, GR-70013 Heraklion, Crete, Greece.

P.Coveney(p.v.coveney@ucl.ac.uk), S.Wan (shunzhou.wan@ucl.ac.uk), S.Manos (s.manos@ucl.ac.uk) and S. Zasada (stefan.zasada@ucl.ac.uk) are with the University College London.

A. Folarin (a.folarin@ucl.ac.uk) is with the Cancer Institute, University College London, London, UK.

P.Büchler (Philippe.Buechler@artorg.unibe.ch), T.Bardyn (thibaut.bardyn@artorg.unibe.ch), S. Bauer (stefan.bauer@istb.unibe.ch), M. Reyes (mauricio.reyes@istb.unibe.ch).

T.Bily(tomby@kam.mff.cuni.cz), V.Bednar (bednar@kam.mff.cuni.cz), M. Karasek (michal.karasek@gmail.com) are with Charles University in Prague Faculty Mathematics and Physics, Department of Applied Mathematics, Prague 1, 118 00, Czech Republic.

N. Graf (Norbert.Graf@uniklinikum-saarland.de) is with the University of the Saarland, Department of Pediatric Oncology, 66421 Homburg, Germany.

R. M. Bohle (rainer.bohle@uniklinikum-saarland.de) and Y.-J- Kim (yoo-jin.kim@uniklinikum-saarland.de) are with the University of the Saarland, Department of Pathology, 66421 Homburg, Germany.

E. Meese (hgemee@uniklinik-saarland.de) is with the University of the Saarland, Department of Human Genetics, 66421 Homburg, Germany

H. Stenzhorn (holger.stenzhorn@ifomis.uni-saarland.de) is with IFOMIS, University of Saarland.

G.Clapworthy(gordon.clapworthy@beds.ac.uk), E.Liu (enjie.liu@beds.ac.uk) are with the University of Bedfordshire, Luton, LU1 3JU, UK

J. Sabczynski (Joerg.Sabczynski@philips.com) is with Philips Technologie GmbH Forschungslaboratorien, Hamburg, Germany.

I. Introduction

THE accelerated accumulation of important experimental and clinical data as well as biological and medical knowledge pertaining to the *natural phenomenon* of cancer renders quantitative integration over many scales of biocomplexity a pressing necessity. Specially designed mathematical and computational models and systems appear to constitute the only rational way to respond to this need. The long term goal of model and related system development is to optimize treatment in the patient individualized context by exploiting quantitative understanding of cancer in conjunction with the exploitation of available multiscale biodata. Obviously, a strict and lengthy clinical adaptation and validation process is an inelastic prerequisite for the eventual clinical translation of such models and systems. Within this framework the European Commission funded project ContraCancrum (contracancrum.eu) includes the development of an integrated multiscale malignant tumor simulator (*Oncosimulator*) able to simulate tumor growth and response to treatment across biocomplexity scales. Glioblastoma multiforme and lung cancer are addressed as two tumor paradigms.

Regarding the mathematical approaches adopted on the cellular and higher biocomplexity levels the following decisions have been taken. Diffusive tumor growth is addressed primarily via a diffusion equation based

continuous description which is transformed into an algebraically solvable finite difference or finite element formulation. On the other hand tumor response to treatment is addressed primarily via a discrete entity – discrete event formulation.

This document provides a brief outline of the *Oncosimulator* components and the integration policy adopted. It also briefly outlines an elementary problem addressed which is nevertheless indicative of the many stages of model development. The short communication concludes with a brief discussion on the characteristics of the major modeling approaches that refer to the cellular and higher biocomplexity levels since the latter constitute the basis for the entire Oncosimulator integration.

II. A Brief Outline of the Oncosimulator Component Integration Strategy

The following indicative biomodels are currently under integration yet at a varying level of mutual linkage:

i. Patient specific chemotherapy drug targeting (e.g. interactions between ligand/inhibitor and receptor tyrosine kinases in gliomas and lung cancer). Molecular dynamics simulations.
ii. Molecular interdependence networks for cell survival probabilities
iii. Non imageable tumor growth
iv. Angiogenesis
v. Invasion via cell diffusion
vi. Imageable tumor growth
vii. Tumor response to chemotherapy, radiotherapy and radio-chemotherapy
viii. Tumor and normal tissue biomechanics
ix. Macroscopic image analysis of the anatomic region of interest
x. Consideration of treatment limitations imposed by normal tissue toxicities induced by candidate treatment schemes and / or schedules.

Elementary model integration is heavily based on the *top-down*, discrete entity – discrete event multiscale cancer modeling approach [1-10] since it provides great flexibility which is crucial to achieve that goal .

The adopted sequence of the major integration steps is the following:

a. Fusion of the **T**umor growth and treatment response models with the respective **B**iomechanical models to produce the *TB integrator*
b. Fusion of the *TB integrator* with the **I**mage analysis module to produce the *TBI integrator*
c. Fusion of the affected **N**ormal tissue module with the *TBI integrator* to produce the *TBIN integrator*
d. Fusion of the **M**olecular simulations and networks module with the *TN integrator* to produce the *TBINM integrator*
e. Logical and technical testing and optimization of the integrated simulation system before undergoing clinical adaptation and validation.

In order to integrate biological mechanisms acting on different biocomplexity scales (levels) the *summarize and jump* strategy delineated in [9] has been adopted.

III. Introduction to One of the Mathematical Approaches Adopted

As an example of the mathematical approaches adopted the diffusion equation based treatment is briefly outlined. In the continuum context tumor growth can be approximately expressed by the following differential equation [11,12] formulated in words as:

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

and the appropriate boundary conditions.

Using standard mathematical symbols the previous statement takes the following form:

$$\left\{\begin{array}{l} \dfrac{\partial c}{\partial t} = \nabla (D \nabla c) + \rho c - G(t)c \quad in\ R \\ c(\mathbf{x},0) = f(\mathbf{x}), \quad initial\ condition \\ \mathbf{n} \cdot D \nabla c = 0 \quad on \quad \partial R, \quad boundary\ condition \end{array}\right\} \tag{1}$$

The variable c denotes the cell concentration at any spatial point and time t. The parameter D denotes the diffusion coefficient and represents the active motility of tumor cells. The term ρ denotes the net rate of tumor growth including proliferation, loss and death. $G(t)$ accounts for the temporal profile of treatment such as radiotherapy and/or chemotherapy. $\mathbf{n}$ is the unit normal to the boundary ∂R of the domain R and f is a known function. In the mathematical model implemented as a first approximation $G(t)=k$ is constant. To solve the above equation and boundary condition system several methods such as the Crank Nicholson technique in conjunction with the conjugate gradient solver have been applied.

IV. Indicative Intermediate Investigational Results

The following simple problem can be viewed as a typical example of the many developmental and exploratory stages of the *Oncosimulator* construction. In order to check an implementation of the Crank-Nicholson numerical scheme used in conjunction with the conjugate gradient method for the solution of the diffusion equation for glioma growth [11,12], a simple *artificial* problem has been defined and solved. Equation (1) has been solved for the cubic region (7 cm x 7 cm x 7 cm, x = 7 cm) with diffusion coefficient D = 0.0065 cm^2/day, time step dt = 1 sec, dx = 0.1 cm,

growth rate = 0.012 units/day and loss rate = 0.0013 units/day. The initial shape of the tumor has been a sphere of radius=0.2 cm and cell density 35000 tumor cells per mm^3. Cell density falls abruptly to zero outside the sphere (which is a biologically unrealistic scenario). Skull bone has been assumed to lie on the boundaries of the cubic mesh (boundary conditions). Skull bone does not practically allow glioma cells to be diffused through it. Execution of the simulation code for a time interval equal to 90 days has been performed. The tumor cell concentration as a function of time is presented on day 45 and day 90 (Fig.1). The predictions are in agreement with the diffusion aspects of the problem provided that the rate values are taken into account. It should be stressed, however, that the artificial problem outlined so far has been designed only to facilitate mathematical validation of the corresponding code and by no means as a representation of the complete glioblastoma invasion problem.

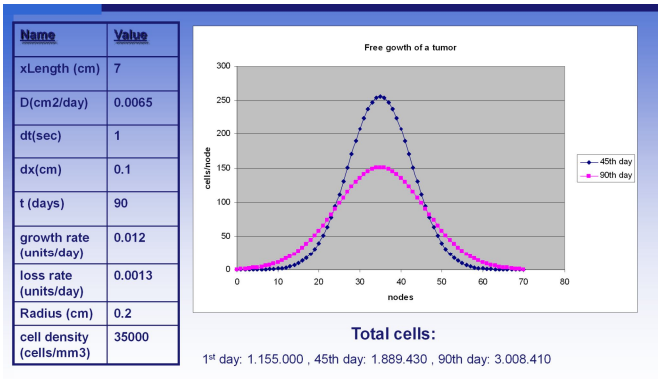


Fig.1 Tumor cell concentration (initial spherical tumor cell density = 35000 cells/ mm^3)

V. DISCUSSION

Due to the wide scope of the ContraCancrum *Oncosimulator* system, several modeling approaches have been adopted. Referring to the cellular and higher biocomplexity levels which provide the matrix for the overall *Oncosimulator* integration, two major *cancer modeling schools* can be identified. The *continuous entity based formulation school* (having the diffusion equation at its core (in practice in a *finitized* form)) and the *discrete entity - discrete event formulation school* (dealing with purely discrete notions such as proliferative potential cell categories, cell cycle phases, cycling or non cycling cell states etc.). Both intuition and research experience suggest that none of the formulations can perform equally efficiently when dealing with (a) markedly diffusive tumor growth and (b) tumor response to treatment. The continuous entity based formulation has shown to be better positioned in answering questions of the type "*what is the real spatial extent and the actual concentration profile of a glioblastoma tumor within the brain (including both imageable and non imageable components) ?*" On the other hand the discrete entity – discrete event formulation has shown to be better positioned in answering questions of the type "*what will spatiotemporally be the biological constitution of a tomograhically, histopathologically and molecularly characterized tumor following administration of a chemotherapeutic cycle of a given drug?*". Therefore, care should be taken in formulating the questions to be addressed by mathematical and computer modeling before proceeding to the formulation of the modeling approach.

REFERENCES

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

[2] Dionysiou D.D., 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 , 1-20 , 2004

[3] G.S.Stamatakos, D.D.Dionysiou, N.M.Graf, N.A.Sofra, C.Desmedt, A.Hoppe, N.Uzunoglu, M.Tsiknakis , 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 , Proceedings of the 29th Annual International Conference of the IEEE EMBS Cite Internationale, August 23-26, SuB07.1: 6628-6631 , Lyon, France , 2007

[4] Dionysiou, D.D., G.S. Stamatakos, D. Gintides, N. Uzunoglu K. Kyriaki. 2008. Critical parameters determining standard radiotherapy treatment outcome for glioblastoma multiforme: a computer simulation. Open Biomed Eng J 2: 43-51.

[5] Graf, N., A. Hoppe , E. Georgiadi, R. Belleman, C. Desmedt, D. Dionysiou, M. Erdt , J. Jacques, E. Kolokotroni, A. Lunzer, M. Tsiknakis and G. Stamatakos. 2009. "*In silico* oncology" for clinical decision making in the context of nephroblastoma. Klin Paediatr 221: 141-149.

[6] Kolokotroni E.A.,, G. S. Stamatakos, D. D. Dionysiou, E. Ch. Georgiadi, Ch. Desmedt, N. M. Graf, "Translating Multiscale Cancer Models into Clinical Trials: Simulating Breast Cancer Tumor Dynamics within the Framework of the "Trial of Principle" Clinical Trial and the ACGT Project.," Proc. 8th IEEE International Conference on Bioinformatics and Bioengineering (BIBE 2008), Athens, Greece, 8-10 Oct. 2008. IEEE Catalog Number: CFP08266, ISBN: 978-1-4244-2845-8, Library of Congress: 2008907441, Paper No. BE-2.1.1, length: 8 pages (in electronic format). 2008

[7] Georgiadi E. Ch., G. S. Stamatakos, N. M. Graf, E. A. Kolokotroni, D. D. Dionysiou, A. Hoppe, N. K. Uzunoglu, "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," Proc. 8th IEEE International Conference on Bioinformatics and Bioengineering (BIBE 2008), Athens, Greece, 8-10 Oct. 2008. IEEE Catalog Number: CFP08266, ISBN: 978-1-4244-2845-8, Library of Congress: 2008907441, Paper No. BE-2.1.2, length: 8 pages (in electronic format). 2008

[8] Giatili S.G., G. S. Stamatakos, D. D. Dionysiou, E. A. Kolokotroni, E. Ch. Georgiadi, Geometrical and Mechanical Aspects of Tumor Growth and Response to Chemotherapeutic Schemes in the Context of the ACGT Oncosimulator," Proceedings of the 3rd International Advanced Research Workshop on In Silico Oncology, Istanbul, Turkey, Sept. 23-24, 2008. Edited by G. Stamatakos and D. Dionysiou pp.35-37, 2008

[9] Stamatakos G. and D. Dionysiou. 2009. Introduction of hypermatrix and operator notation into a discrete mathematics simulation model of

malignant tumor response to therapeutic schemes *in vivo*. Some operator properties. Cancer Informatics 2009:7 239–251.

[10] Stamatakos, G.S., E. A. Kolokotroni, D. D Dionysiou, E. C Georgiadi, C. Desmedt, An advanced discrete state - discrete event multiscale simulation model of the response of a solid tumor to chemotherapy. Mimicking a clinical study. J.Theor. Biol. *In press* (2010).

[11] Murray J.D. Mathematical Biology II: Spatial Models and Biomedical Applications. Third Edition, Springer-Verlag, 2002; Chapter 11

[12] Swanson K. R., Carly Bridge, J.D. Murray, Ellsworth C. Alvord Jr. Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. Journal of the Neurological Sciences 2003; 216:1-10.

V. Biomechanics in Tumor Modeling

Biomechanical model of tumor growth: application to the European ContraCancrum project

Christian May, Thibaut Bardyn, Mauricio Reyes, Philippe Büchler

***Abstract*—This abstract presents the biomechanical model that is used in the European ContraCancrum project, aiming at simulating tumor evolution in the brain and lung. The construction of the finite element model as well as a simulation of tumor growth are shown. The construction of the mesh is fully automatic and is therefore compatible with a clinical application. This biomechanical model will be later combined to a cellular level simulator also developed in the project.**

I. Introduction

THE European ContraCancrum project is developing a composite multilevel platform for simulating malignant tumor development and tissue response to therapeutic modalities and treatment schedules in the brain and lung. The oncosimulator that will be developed in the framework of this project combines a cellular [1] and a biomechanical model. This biomechanical model aims at calculating the mechanical state in the anatomy around the tumor during its growth. This mechanical information is then transferred to the cellular simulator which will predict the behavior of the cells within this tumor. The first challenge is to develop a fully automatic meshing algorithm for the different tissues considered. Based on this mesh, a continuum finite element model is proposed to simulate the tumor, its growth and the mechanical perturbations induced on the surrounding healthy tissues. This abstract presents the algorithm used to create this mesh. A finite element simulation of tumor growth combining reaction/diffusion and mass effect is then performed with this mesh. Examples are given on a model of the brain.

Christian May is with the Computational Bioengineering Group, ISTB Research Center, Stauffacherstrasse 78,3014 Bern, Switzerland (e-mail: christian.may@istb.unibe.ch).

Thibaut Bardyn is with the Computational Bioengineering Group, ISTB Research Center, Stauffacherstrasse 78,3014 Bern, Switzerland (e-mail: thibaut.bardyn@istb.unibe.ch).

Mauricio Reyes is with the Medical Image Analysis Group, ISTB Research Center, Stauffacherstrasse 78,3014 Bern, Switzerland (e-mail: mauricio.reyes@istb.unibe.ch).

Philippe Büchler is with the Computational Bioengineering Group, ISTB Research Center, Stauffacherstrasse 78,3014 Bern, Switzerland (e-mail: philippe.buechler@istb.unibe.ch).

II. Methods

A. Creation of the Mesh

The mesh of the brain was constructed from an image atlas with a spherical tumor artificially added. The voxel-based method approach was used. In brief, it consisted in associating a voxel representing tissue with a cubic element in the mesh. This method is particularly fast and fully automatic. However, the surfaces generated are composed of jagged meshes which can produce stress concentrations. One solution to improve the aspect of these meshes is to smooth its external surface.

The outer surface of the mesh was extracted and smoothed according to a geometric signal processing approach [2]. First a frequency decomposition of the surface was done. Then, a low pass filtering was applied to the decomposition in order to remove high frequencies (sharp edges). The cutoff frequency k of the filter defined the degree of smoothing. This algorithm is numerically stable, and preserves the volume. Smoothing was performed up to a frequency of k=0.03. Extensive smoothing created distorted elements with large or very small dihedral angles on the surface of the mesh. To improve the quality of the mesh, hexahedral elements bearing a large angle between faces were divided into prism elements. If the angle was superior to 130 °, then the edge at the intersection of the two faces was used for the division. The element was divided by the "virtual" plane that joined this edge with its opposite in the element.

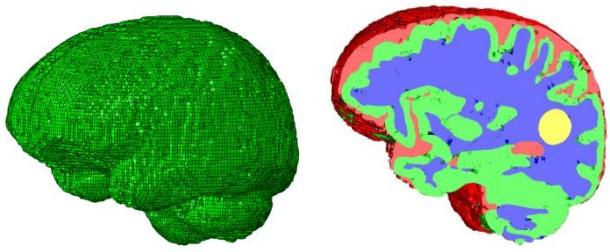


Fig. 1. Finite element model of the brain used for the simulation of the tumor growth. The different colors represent the anatomical structures with distinct material properties

B. Growth of the tumor

A coupling approach combining of diffusion/reaction and mass effect [3] was simulated in the open source finite element code FEBio [4].

A diffusion/reaction step was used to simulate the invasion of cancerous cells in the healthy tissue. A mass effect simulation (increase in volume of the tumor) was then performed to grow the elements proportionally to the quantity of cancerous cells that had been produced and diffused in each element of the mesh. The deformation of the mesh was then used to initialize a new diffusion step. The process looped by alternating between the two simulation types and was solved sequentially. In the future the reaction/cell production part of this simulation will be provided by the cellular simulator developed in the ContraCancrum project [1]

The mass effect was considered as isotropic and uniform. The change in volume was modeled as a uniform strain added in the three main directions in the elastic formulation of the element. A linear relationship was assumed between cell concentration increase and volume change. If the density of cells reached or was equal to the maximum carrying capacity of the tissue ($3.5E^4$ Cells.mm^{-3}) then a fixed growth strain was imposed to the element (a value of 2% was taken arbitrarily). Otherwise, the increase of volume depended linearly on the number of cancerous cells. At each step, the displacement at every node was transmitted to the diffusion simulation. Similar material properties were used for grey and white matter (E=694Pa, ν=0.3). The experimental data on the mechanical properties of tumor is extremely scarce. Therefore, as suggested by Wittek et al. [5], it was assumed that the properties of the tumor were close to those of the surrounding soft tissue. Considering the fact that the simulation was performed for a large time, viscoelastic effects were neglected and finite elastic properties were used. The skull was not included in the model. Instead, displacement of the nodes on the external surface of the model was fully constrained.

The diffusive part of tumor growth was simulated using the reaction diffusion law:

$$\frac{\partial c}{\partial t} = div\left(D_t \nabla c\right) + \rho c \qquad (1)$$

Where c is the concentration of cancerous cells, D_t is the diffusion tensor indicating the preferential direction of diffusion in the tissue. For this example, the tissues were considered as isotropic. The diffusivity properties of the different tissues were defined so that the white matter diffusivity ($86400E^{-5}$ mm^2.day^{-1}) was a hundred times larger than that of the grey matter ($86400E^{-7}$mm^2.day^{-1}). The cerebro spinal fluid was considered as a no diffusive material. The boundaries of the models were defined so that no diffusion was possible outside of the brain tissue (Neumann boundary condition). The original tumor was constrained so as to have a constant concentration of cells ($3.5E^4$ Cells.mm^{-3}). The source parameter ρ representing the production of cancerous cells per day was set to $2.2E^{-5}$day^{-1}.

III. RESULTS

The brain mesh consisted of 255841 nodes and 194582 elements among which 94215 hexahedral elements and 161626 prism elements (figure 1). The whole process (including generation of the mesh, smoothing, correction via prism division and generation of the XML file) took 3.28 minutes on a Pentium 4/2.4GHz CPU with 2GB of RAM and was fully automatic.

The tumor growth simulations were performed for a virtual period of six months (figure 2). The results showed that growth primarily occurred in tissues where the diffusivity was higher (white matter) and in the tumor. The structures in white matter were obviously enlarged after the simulation time. The lack of diffusion in the CSF and the grey matter was clearly visible on the color plots. The large span of cancerous cells in the brain tissue proved that the diffusion component has to be considered for the simulation of the growth of the tumor.

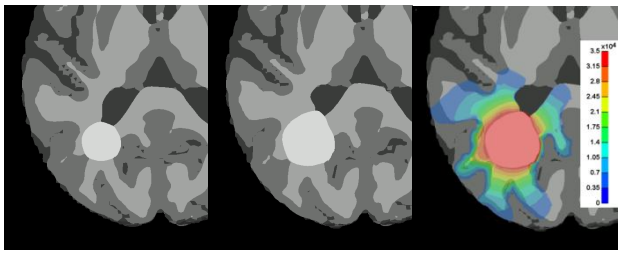


Fig. 2. Result of the coupled growth for a slice of the brain MRI. From left to right, original configuration of the brain, deformed configuration of the brain showing the mass effect of tumor growth, result of the reaction/diffusion simulation mapped onto the deformed model

IV. CONCLUSIONS

The steps constituting the biomechanical model in the oncosimulator were presented in this paper. The creation of the mesh was made as fast and automatic as possible to fulfill the requirements of a clinical application. The voxel mesh approach, which is extremely efficient in terms of time, has been enhanced with a smoothing algorithm which improves its accuracy. Accurate stress information is a priority when designing the model since simulation on cell level relies on the calculated pressure in the tissues.

The mesh was tested with a coupled mass effect and diffusion/reaction model. Results show selected growth of the tumor in tissues where diffusivity is higher. This model proves the feasibility of this type of analysis and is ready to be integrated in the general oncosimulator of the project.

REFERENCES

[1] G. S. Stamatakos, V. P. Antipas and N. K. Uzunoglu, "A spatiotemporal, patient individualized simulation model of solid tumor response to chemotherapy in vivo: the paradigm of glioblastoma

multiforme treated by temozolomide," *IEEE Trans. Biomed. Engineering*, vol. 53, No. 8, pp. 1467-1477, 2006.

[2] G. Taubin, T. Zhang and G. Golub, "Optimal surface smoothing as filter design," *Proceedings of the 4th European Conference on Computer Vision*, vol. 1064, pp. 283-292, 1996

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

[4] S. A. Maas, B. J. Ellis, D. S. Rawlins, and J. A. Weiss, "A comparison of FEBio, ABAQUS, and NIKE3D. Results for a suite of verification problems," *UUSCI-2009-009*, Scientific Computing and Imaging Institute, University of Utah, Salt Lake City, UT 84112 USA, 2009

A. Wittek,K. Miller, R. Kikinis and S. K. Warfield, "Patient-specific model of brain deformation: application to medical image registration," *J.Biomech, vol.*. 40**,** pp. 919-929, 2007.

A Markov-Random-Field-Based Biomechanical Tumor Growth Model for Atlas-Based Segmentation of Brain Tumor Images

Stefan Bauer, Mauricio Reyes

***Abstract*—We show the application of a new technique for soft tissue deformations to brain atlas images. The proposed method consists of computing local voxel displacements based on a Markov Random Field approach, taking into account biomechanical tissue properties. The technique is designed for atlas-based segmentation of brain tumor images within a clinically-oriented workflow. It offers the possibility to modify a healthy brain atlas by introducing a tumor seed and grow the tumor to its approximate patient tumor shape, simulating mass effect on the surrounding tissues. Subsequently the patient image can be implicitly segmented by registering the modified tumor-bearing brain atlas to the pathologic patient image.**

I. Introduction

FOR image analysis and biomechanical modeling the accurate segmentation of important brain structures is of significant interest. An established way to classify tissue types in healthy humans is to do atlas-based segmentation of different tissue types like grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) [1]. However, in case of brain tumor images this approach fails due to the missing tumor prior in the atlas image. Solutions to this problem have been suggested by several groups [2-4]. The common idea is to introduce a tumor seed into the atlas and grow it to its approximate patient-specific shape using different methods. Cuadra et al. [2] use a simple model of lesion growth without any biomechanical justification, while Mohamed et al. [3] use a more complex finite element model (FEM) considering mechanical tissue properties. Zacharaki et al. [4] employ a linear elasticity model on regular grids. A concise overview of the state of the art was presented by Angelini et al. in [5].

FEM-based methods offer the desired possibility to incorporate mechanical properties into the tumor-induced deformations. However, they suffer from the need to create a volumetric mesh and usually require a significant amount of computation time. Furthermore, the mesh-free method suggested in [4] runs on a subsampled version of the image in order to circumvent the high computational cost.

Manuscript received May 5, 2010.

Stefan Bauer and Mauricio Reyes are with the Institute for Surgical Technology and Biomechanics, University of Bern, Switzerland (phone: +41 31 631 5948, fax: +41 631 5960, e-mail: stefan.bauer@istb.unibe.ch).

II. Materials and Methods

We suggest a clinically-oriented and mesh-free method to model soft-tissue deformations, which is applicable to atlas-based segmentation of brain tumor images. The deformation technique is based on finite differences in a local neighborhood of each voxel using Markov Random Fields (MRF). The idea was initially proposed by Seiler et al. in [6] and applied to 2-dimensional atlas-based segmentation of brain tumor images in [7]. Here we show the extension of the tumor growth model to 3D as well as preliminary results on patient data.

A. Hierarchical Displacement Model

The general idea to model soft tissue deformations is to minimize an energy function of the type

$$U_{total} = U_{prior} + U_{observation} \quad (1)$$

U_{prior} represents the biomechanical information of the brain tissues and $U_{observation}$ introduces boundary conditions. Minimization of these energies is done in cliques of a neighborhood system surrounding a center voxel. The local tissue characteristics are based on their Young's modulus. The minimization of the chosen energy function yields the tumor-induced growth and deformation according to the given tissue characteristics and boundary conditions.

The displacement model is applied to the atlas image in a hierarchical way, which offers the possibility to use fast and stable local optimizers. As a local optimizer, iterative conditional modes (ICM) is used to find the best solution at each level of hierarchy.

B. Application to Tumor Growth Modeling

We assume a radial, outward pushing force of the tumor. The growth is performed in an iterative way in order to circumvent warping problems. When the desired patient tumor shape is attained, tumor growth is stopped. We want

to emphasize that this deformation method is not intended to be a viable tumor growth model, but should be rather considered as a fast and simple, yet biomechanically justified technique to introduce tumor-induced deformations into an atlas image. This modified atlas image later serves as input for non-rigid registration to the pathologic patient image in order to recover the remaining deformations and thus, implicitly segment the different tissue types.

III. Results

As a reference to demonstrate the modeling results, we chose the freely available SRI24 atlas [8]. This average atlas is well-suited for the intended registration purpose thanks to its sharpness.

In figure 1 a 3D volume rendered image of the modified atlas label map with tumor is shown. The initial and the final step of the tumor growth and tissue displacement process are depicted. At the top, a rendering of the atlas with the initial tumor seed is shown. For better visualization it is cut open at the tumor seed location. The picture in the middle shows the same atlas rendering at the final iteration of the displacement process. It can be observed how the surrounding tissues are deformed. Despite the radially outward pushing tumor, the deformation is not exactly circular due to the different tissue characteristics as can be seen from the axial zoom image in the middle row of figure 1. The bottom part of figure 1 presents the magnitude of the displacement vector field (DVF) applied to each voxel in the 3D image. The displacement is largest inside the tumor area, but also the surrounding tissue are affected. The impact of the tissue displacement is decreasing, the further the voxels are away from the current tumor region.

The 4 different labels for WM, GM, CSF and tumor can be used to segment the patient image after application of a non-rigid registration technique.

IV. Discussion and Conclusion

A simple and clinically-oriented method to deform brain tissues in an atlas image was presented. This method can be used as an initialization step to modify a healthy brain atlas in order to be able to perform atlas-based segmentation of brain tumor images. The atlas labels can be warped to the patient image with the deformation field, which is obtained using non-rigid registration techniques for matching the modified atlas to the pathologic patient image.

MRF approaches are well-suited for parallelization which offers the possibility to significantly speed up the deformation process in the future by exploiting GPU-based computations.

Acknowledgment

Funding by the European Union within the framework of the ContraCancrum project (FP7 -- IST-223979) is gratefully acknowledged.

References

[1] D. Collins C. Holmes, T. Peters, A. Evans “Automatic 3D model-based neuroanatomical segmentation” Human Brain Mapping, vol. 3, 1995

[2] M. Cuadra, C. Pollo, A. Bardera, O. Cuisenaire, J. Villemure, J. Thiran “Atlas-based segmentation of pathological MR brain images using a model of lesion-growth” IEEE Transactions on Medical Imaging, vol. 23, 2004

[3] A. Mohamed, E. Zacharaki, D. Shen, C. Davatzikos “Deformable registration of brain tumor images via a statistical model of tumor-induced deformation”, Medical Image Analysis, vol. 10

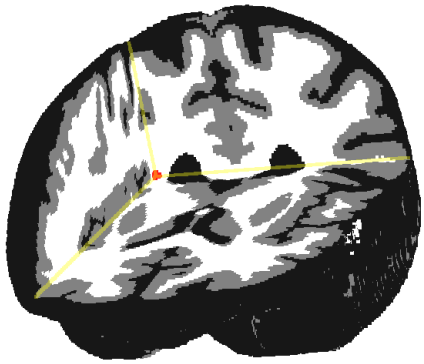
[4] E. Zacharaki, C. Hogea, D. Shen, G. Biros, C. Davatzikos “Non-diffeomorphic registration of brain tumor images by simulating tissue loss and tumor growth” NeuroImage, vol. 46, 2009

[5] E. Angelini, O. Clatz, E. Mandonnet, E. Konukoglu, L. Capelle, H. Duffau “Glioma Dynamics and Computational Models: A review of segmentation, registration and in silico growth algorithms and their clinical applications”, Current Medical Imaging Reviews, vol.3, 2007

[6] C. Seiler, P. Buechler, L.P. Nolte, M. Reyes, R. Paulsen “Hierarchical Markov Random Fields applied to model soft tissue deformations on graphics hardware”, Recent Advances in the 3D Physiological Human, 2009

[7] S. Bauer, T. Bardyn, P. Buechler, M. Reyes “Atlas-Based Segmentation of Brain Tumor Images using a Markov Random Field-Based Tumor Growth Model and Non-Rigid Registration”, submitted

[8] T. Rohlfing, N. Zahr, E. Sullivan, A. Pfefferbaum “The SRI24 multi-channel brain atlas : construction and applications”, Proceedings of the Society of Photo-Optical Instrumentation Engineers, vol. 6914, 2009



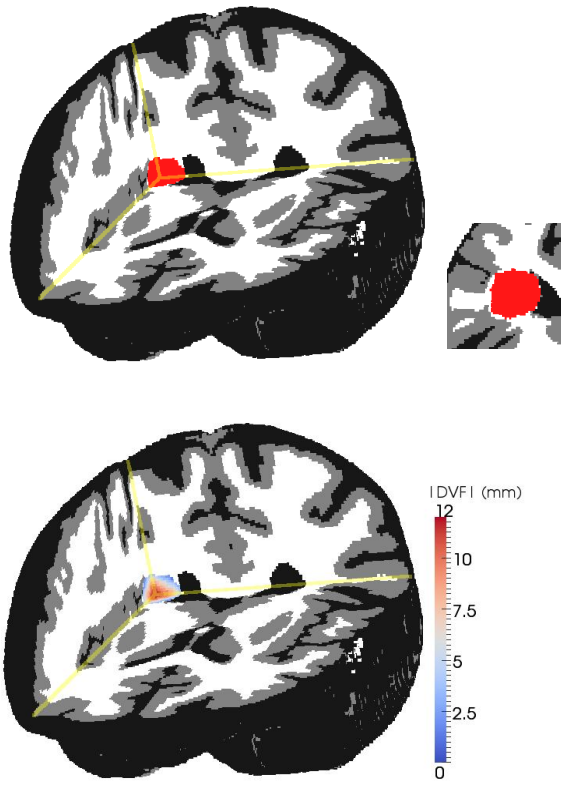


Fig. 1. 3D volume rendering of the modified SRI24 atlas. CSF is black, grey matter grey, white matter white and the tumor is drawn in red. From top to bottom: Atlas with tumor seed, modified atlas after 6th growth iteration and zoom on the tumor region in axial view, magnitude of the DVF at each voxel position.

VI. Continuous Formulation Based Tumor Modeling

G. Stamatakos and D. Dionysiou (Eds): *Proc. 4th Int. Adv. Res. Workshop on In Silico Oncology and Cancer Investigation (4th IARWISOCI) – The ContraCancrum Workshop*, Athens, Greece, Sept. 8-9, 2010 (www.4th-iarwisoci.iccs.ntua.gr)

Approximating the diffusion – reaction equation for developing glioma models for the ContraCancrum Project: a showcase

Alexandros Roniotis, Konstantinos Panourgias, John Ekaterinaris, Kostas Marias, Vangelis Sakkalis

***Abstract* – During the last decades, predictive oncology has pursued novel methods for simulating the mechanism that govern the behavior of tumors. Especially for the case of glioblastoma, several models have been developed towards coupling diffusive behavior with cell proliferation. A glioma model could ideally be used by clinicians for predicting the spatiotemporal evolution of tumor and the response to different therapeutic treatments. The most common models use the diffusion – reaction equation for simulating the change of glioma cell density in space and in time. Such a model has been developed for the needs of the ContraCancrum project and has been described in previous papers. This paper presents how a newly developed 3D Finite Elements framework can be applied for the glioma growth simulation. Moreover, we present some results on applying this model on a controllable test case.**

I. Introduction

GLIOBLASTOMA multiforme is the most malignant form of brain tumor, characterized as a WHO-4 glioma, with average remaining life being at 12 months after detection. Glioma exhibits a highly diffusive behavior, making the detection of the exact tumor boundaries with common imaging techniques (PET, MRI, CT) enigmatic. Thus, glioblastomas and other milder types of glioma have been extensively studied during the last two decades by researchers. However, mortality scores still remain high, raising the interest for developing mathematical models that can approximate the behavior of glioma. Such models aim at coupling the highly invasive behavior of gliomas with the proliferation of tumor cells. Therefore, variations of the diffusion – reaction equation have been proposed towards this direction. Hence, the models that have been developed using this equation are called *diffusive*.

Manuscript received May 5, 2010. Acknowledge support from EC Contra Cancrum Project ICT-2007-223979

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

K. Panourgias and J. Ekaterinaris are with the Institute of Applied & Computational Mathematics at FORTH, Vassilika Vouton, GR-70013 Heraklion, Crete, Greece ({panourgias; ekaterin}@iacm.forth.gr).

II. Background

The initial diffusive models [1] used a simplified version of the diffusion equation, for simulating the spatiotemporal change of glioma cell density. The equation that was used is:

$$\frac{\partial c}{\partial t} = div(D\nabla c) + f(c)$$

where $c(\mathbf{x},t)$ is the tumor concentration at time t and in position $\mathbf{x} = (x, y, z)$, D is the diffusion coefficient, ∇ and *div* are the gradient and divergence operators respectively and $f(c)$ is the net cell proliferation rate. $f(c)$ can be adjusted to simulate different proliferative schemes, such as geometrical, Verhulst, Gompertz and chemotherapy.

This primitive model was extended by Swanson [2], in order to incorporate the heterogeneous motility of glioma cells in white and gray matter. Indeed, there is a fivefold difference between the speed of cell migration in gray and white matter. Thus, the heterogeneous model used:

$$\frac{\partial c}{\partial t} = div(D(\mathbf{x})\nabla c) + f(c)$$

where $D(\mathbf{x})$ is the local diffusion coefficient, that is D_g or D_w if $\mathbf{x}$ is in gray and white matter, respectively and $D_w = 5D_g$.

Later, this model was further extended to incorporate the anisotropic migration of tumor cells, which was observed to be facilitated along white fibers [3,4]. The new model uses a 3x3 diffusion tensor $\mathbf{D}(\mathbf{x})$, instead of $D(\mathbf{x})$. We have used this last model version for simulating tumor growth for the needs of the ContraCancrum Project [5], by using Finite Elements methods. Diffusive models that use Finite Differences have already been developed.[6].

III. The Model

A. Con traCancrum Project

ContraCancrum project aims to pave the way for translating clinically validated multilevel cancer models into clinical practice. The models will assist the clinician to define the optimal therapy for the individual patient taking into consideration all the available clinical data from different scales (molecular to tissue level), modalities (e.g. multimodal cancerimaging) and examinations (e.g. before and after therapy).

B. Glioma Modeling Test Case

One important component of the ContraCancrum Project is the implementation of glioma diffusive models. One such model has been developed for approximating the solution of the 3d diffusion – reaction equation. In order to make an initial estimation of the model accuracy we have used the model for a simplified test case for which there is known algebraic solution. However, this model is designed for implementing all previously referred models [1-4] and can be adjusted to incorporate different proliferative schemes for untreated [7] and treated tumors [8].

C. Results

For the needs of demonstration, Fig.1 presents the results of applying the model on the virtual brain, presented earlier. These results have been obtained by getting the middle plane-intersection with the 3D space. We can observe that the cells are initially concentrated in the centre of the sphere. As time elapses, they are diffused inside the spherical "brain" and the concentration changes become smoother. Indeed, this is expected, because no proliferation term has been used for the needs of the experiment. This model is currently under validation and is applied on real medical images that have been provided for the ContraCancrum Project.

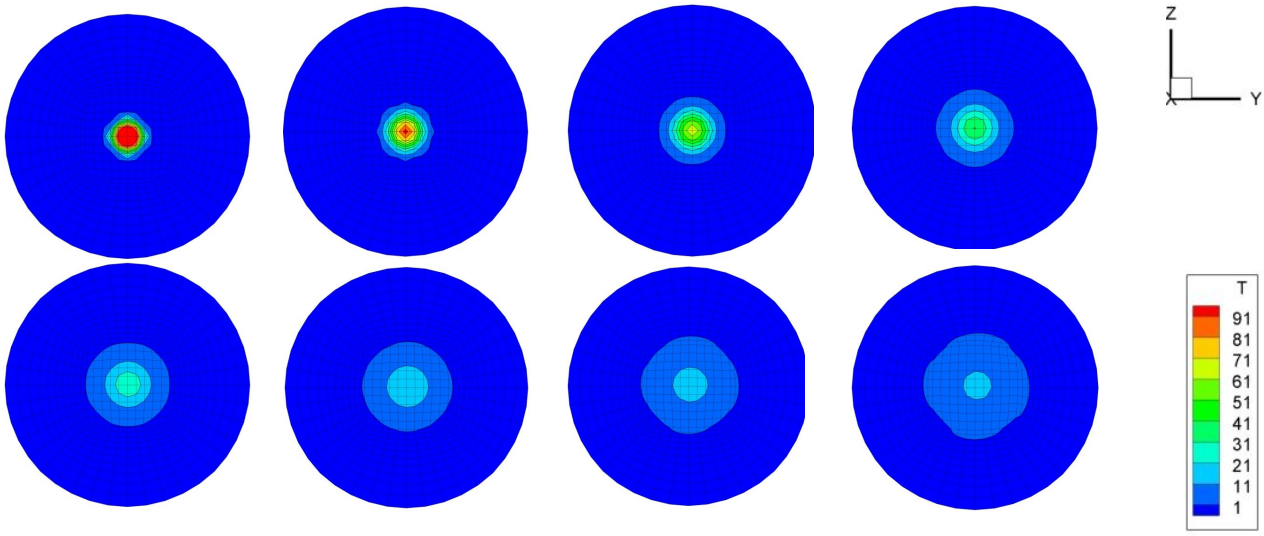


Fig. 1 – Eight different snapshots of the approximated glioma cell concentration taken at iterations t=15, 50, 100, 200, 300, 400, 500 and 575

The model uses the Local Discontinuous Galerkin method which is a finite element method. According to this, each element has its own interpolated function and there is discontinuation on the boundaries. Therefore, the solution can be found by iterative proceeding, element by element, till the system converges.

For this specific experiment, elements use modal (hierarchical) basis function with first order polynomial as maximum basis ($p1$) which give order of accuracy $O(2)$. The basis functions are orthogonal Legendre polynomials, which are a subcase of Jacobi polynomials. For time marching, we use explicit Runge-Kutta third order scheme.

The domain of solution is a spherical geometry with radius R=3.46 m and covered by Hexahedral elements. Due to geometry, an interior cubic needed be constructed, filled by hexahedrals.

As initial condition, we set a constant value c_0=100 cells/mm^2 within a spherical domain having radius $r = 0.3$m. Beyond this radius, c_0 was set to 0. On the boundaries, we specified a constant value c=0. The time step for stability was initially set at ΔT=0.00001 and the total number of steps was n=1722. In the latest step the time was 0.017 sec.

IV. Conclusions

This paper presents a simplified test case of a glioma diffusive model, which has been designed for the purposes of the ContraCancrum Project. Unlike previously implemented models [6], here we have developed a 3D Finite Elements framework. This method allows considerable freedom in adding computational elements wherever is necessary. This is important when dealing with highly irregular geometries, such as brain. Moreover, this Finite Elements framework can be also used for simulating the biomechanical deformations caused by the tumor growth and also take into consideration other important factors (e.g. convection). The model is currently being tested on real medical data, provided by the Saarland University, Germany.

References

[1] P. Tracqui, "From passive diffusion to active cellular migration in mathematical models of tumor invasion", *Acta Bibliotheoretica*, vol. 43, pp. 443-464, 1995.

[2] 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.

[3] S. Jbabdi, E. Mandonnet, and H. Duffau, "Simulation of anisotropic growth of low-grade gliomas using diffusion tensor imaging," *Magn Reson Med*, vol. 54, pp. 616-24, 2005.
[4] 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 tumors in MR images coupling diffusion with biomechanical deformation", *IEEE Transactions on Medical Imaging*, vol. 24, pp. 1334-1346, 2005.
[5] ContraCancrum website: http://www.contracancrum.eu/
[6] A. Roniotis, K. Marias, V. Sakkalis, G. Tsibidis, and M. Zervakis, "A Complete Mathematical Study of a 3D Model of Heterogeneous and Anisotropic Glioma Evolution", *IEEE Eng. in Med. and Biol. Soc. Conf.* vol. 1, pp.2807-2810, 2009.
[7] M. Marusic, Z. Bajzer, J. P. Freyer, and S. Vuk-Palovic, "Analysis of growth of multicellular tumor spheroids by mathematical models," *Cell Prolif*, vol. 27, pp. 73-94, 1994.
[8] K. R. Swanson, E. C. Alvord Jr., J. D. Murray, "Quantifying Efficacy of Chemotherapy of Brain Tumors (Gliomas) with Homogeneous and Heterogeneous Drug Delivery," *Acta Biotheoretica*, vol. 50, no. 4, pp. 223-237, 2002

Glioma diffusive modeling: Calculating diffusion coefficients from atlases with proportional tissue information

Alexandros Roniotis, Vangelis Sakkalis, Georgios Stamatakos, Michalis Zervakis, and Kostas Marias

***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. Diffusive models have been used for simulating the spatiotemporal change of glioma cell concentration. The most recent ones take into account tissue heterogeneity and the anisotropic migration of glial cells along white fibers. The main purpose of this paper is to present a novel method for computing the coefficients for the diffusion, taking into account the proportion of white and gray matter in atlases extracted by medical images. Some initial experiments are presented.**

V. Introduction

GLIOMA, especially glioblastomas, is a leading cause of brain cancer fatality, despite the major advances in medicine during the last decades. Gliomas involve highly invasive and neoplastic growth, making the annotation of the exact lesion boundaries enigmatic to radiologists. This has emerged the necessity to describe the mechanism of glioma growth, using mathematical models. The most recent macroscopic models focus on the growth phase and try to predict the behavior of tumor, in terms of the characteristics of the surrounding and underlying tissue. Moreover, they incorporate parameters for proliferation rates and therapy applications. The most dominant models of this type, namely diffusive models, use the diffusion – reaction equation.

VI. Background

The diffusive models simulate the change of glioma concentration in time and in space, by using two main terms. The first term expresses the diffusive behavior of tumor cells and the way they are dispersed in brain, while the second one constitutes the net proliferation rate of glioma cells. The simplest formalism of the equation is the following:

Manuscript received May 5, 2010. Acknowledge support from EC Contra Cancrum Project ICT-2007-223979

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

A. Roniotis and M. Zervakis are with the Dept. of Electronic & Computer Engineering, Technical University of Crete, 73100, Chania, Greece (email: roniotis@ics.forth.gr, michalis@display.tuc.gr).

G. Stamatakos is with the Institute of Communication and Computer Systems, Dept of Electrical and Computer Engineering, Technical University of Athens, Athens, 15780, Greece (email: gestam@central.ntua.gr)

$$\frac{\partial c}{\partial t} = div(D\nabla c) + f(c) \quad (1)$$

where $c(\mathbf{x},t)$ is the tumor concentration in position $\mathbf{x}$ at time t, D is the diffusion coefficient, ∇ and *div* are the gradient and divergence operators respectively and $f(c)$ is the net cell proliferation rate. More recent diffusive models take into account the brain tissue heterogeneity, employing cell invasion, which in white matter is almost 5 times faster than in gray matter [1]. They also use anisotropic diffusion tensors [2] for incorporating the migration of glioma cells in brain that is facilitated along the directions of white fibers. Latest methods simulate the mass deformation, by using biomechanics [3].

By including all these changes, the spatiotemporal diffusion equation that describes glioma growth becomes as follows:

$$\frac{\partial c}{\partial t} = div(\mathbf{D}(\mathbf{x})\nabla c) + f(c) \quad (2)$$

where $\mathbf{D}(\mathbf{x})$ is the diffusion tensor, i.e. a 3x3 symmetric matrix of the form:

$$\mathbf{D}(\mathbf{x}) = \begin{bmatrix} D_x(\mathbf{x}) & 0 & 0 \\ 0 & D_y(\mathbf{x}) & 0 \\ 0 & 0 & D_z(\mathbf{x}) \end{bmatrix} \quad (3)$$

$D_i(\mathbf{x}), i = x, y, z$ is the directional diffusion coefficient, expressing local anisotropy of cell migration. The directional coefficient D_i is the weighted version of the diffusion coefficient proposed in [1], i.e.:

$$D_i(\mathbf{x}) = w_i(\mathbf{x})D(\mathbf{x}) \quad (4)$$

The directional weight $w_i(\mathbf{x}) \in [0,1]$ expresses the anisotropic migration of cells along the direction i (x, y or z axis) in point $\mathbf{x}$, while $D(\mathbf{x})$ is the local diffusion coefficient, expressing the different diffusion velocities in white and gray matter tissues. $D(\mathbf{x})$ equals D_g when point $\mathbf{x}$ is located in gray matter and D_w when in white matter. All recent models use a fivefold difference between these two values, with $D_w = 5D_g$.

Information on tissue matter is generally extracted by normal brain atlases that provide the proportion of white and gray matter in each point of brain. In order to map the atlas with the real medical images, the images are registered to the atlas or conversely. Tissue displacement and deformation due to tumor growth is a factor that registration should take into account.

In our approach we propose a new method for computing the diffusion coefficients extracted by atlases.

We have used the SRI24 Brain Anatomy Atlas, generated by template-free non-rigid registration [4]. The rationale is to propose a probabilistic diffusion coefficients scheme in the diffusion reaction equation instead of the fixed diffusion parameters.

VII. Methods

The SRI24 atlas provides information on the proportion $P_w(\mathbf{x})$ of white matter and the proportion $P_g(\mathbf{x})$ of gray matter in each voxel $\mathbf{x}$. This information can be mapped on real medical images, e.g. MRI, by registration. After having extracted the registered versions of the MRI datasets, the tumor growth can be simulated by solving equation (2)[5].

Up to now, the previous models have ignored tissue information that can be extracted by atlases, and have thresholded the proportional tissue matter values into two discrete values of diffusion coefficients: D_g and D_w for gray and white matter respectively. However, by this method, a voxel that is indexed in the atlas as 51% gray matter and 49% white matter is defined to have diffusion coefficient D_g.

We propose that the proportion of white matter cells should also be taken into account for the model. Thus, we suggest a proportional model where the diffusion coefficient $D(\mathbf{x})$ that is entailed in the diffusion tensor of equation (3) is defined as:

$$D(\mathbf{x}) = P_g(\mathbf{x})\, D_g + P_w(\mathbf{x})\, D_w \quad (5)$$

Moreover, if we apply the 5-fold difference between D_w and D_g, which is proposed in [1], then (5) becomes:

$$D(\mathbf{x}) = (P_g(\mathbf{x}) + 5P_w(\mathbf{x}))\, D_g \quad (6)$$

VIII. Results

We have applied this model on one real temporal glioma case consisting of two sets of 15 MRI slices, taken on two different sessions. The first set was obtained on February 25, 2008, while the second session was on March 18, 2008 (23 days later). The initial and the final images have been delineated by a radiologist from Am. Fleming Hospital, Greece.

For the simulation, we have used two different models, the discrete coefficient model and the proportional model. The model parameters of the simulation were set as: D_g= 0.13 mm^2/ day, initial tumor concentration N_0= 200 cells/mm^2, ρ=0.0012 day^{-1}, diffusion simulation time T=23 days and detection threshold N_T=100 cells/ mm^2.

In Fig. 1, we present the 3d reconstruction of the initial tumor (a) and the final tumor (b). In the last column, we present the simulated final tumor for the proportional model. In order to make an estimate of the accuracy for this experiment, we use the final tumor as ground truth. We then compute the metrics of Jaccard (JC), Dice (DS) and Volume Similarity (VS) metrics [6] defined as:

$$JC = TP/(FP+TP+FN)$$
$$DS = 2TP/(FP+2TP+FN) \quad (7)$$
$$VS = 1-|FP-FN|/(FP+2TP+FN)$$

where TP is the number of tumor voxels belonging to both the ground truth and simulated result, FP is the number of tumor voxels belonging to simulated result but not belonging to ground truth and FN is the number of tumor voxels belonging to ground truth but not belonging to simulated tumor. For the proportional model they are

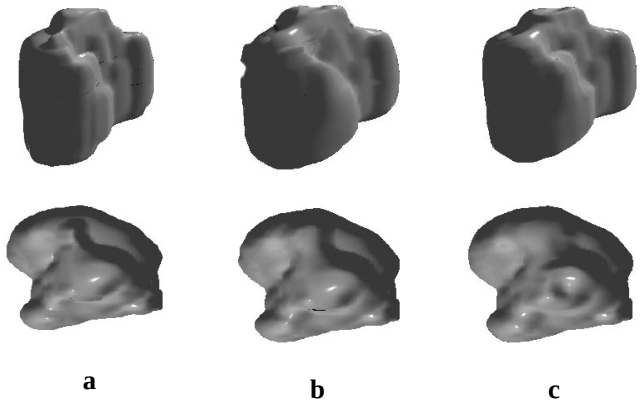


Fig. 1: (a,b) Two different aspects of the 3D reconstruction of the real initial and final tumor. (c) The 3D reconstruction of the simulated tumor, using the proportional model.

JC=94.62%, DS=97.24% and VS=99.39%, while for the discrete coefficient model they are JC=92.01%, DS=94.80% and VS=98.12%.

IX. Conclusions

In this paper we presented an initial evaluation of a diffusion-reaction tumor growth model based on a probabilistic atlas based segmentation of an actual temporal clinical case using a proportional model that takes into consideration partial grey and white matter in each voxel in the diffusion tensor matrix. The results indicate that there is a potential advantage of the probabilistic diffusion coefficients over fixed diffusion parameters. This is an important observation that needs to be validated both from the experimental and the biological interpretation sides. For this reason we are currently working on a large number of glioma cases that have been provided by the Saarland University for the purposes of the ContraCancrum project.

Acknowledgment

The authors would like to thank G. Tsioulos at the Amalia Fleming Hospital, Athens, Greece for annotating the MRIs used in the experiment.

REFERENCES

[1] 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.

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

[3] 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 tumors in MR images coupling diffusion with biomechanical deformation", *IEEE Transactions on Medical Imaging*, vol. 24, pp. 1334-1346, 2005.

[4] T. Rohlfing, N. M. Zahr, E. V. Sullivan, and A. Pfefferbaum, "The SRI24 multichannel atlas of normal adult human brain structure," *Human Brain Mapping*, vol. 31, no. 5, pp. 798-819, 2010.

[5] A. Roniotis, K. Marias, V. Sakkalis, G. Tsibidis, and M. Zervakis, "A Complete Mathematical Study of a 3D Model of Heterogeneous and Anisotropic Glioma Evolution", *IEEE Eng. in Med. and Biol. Soc. Conf.* vol. 1, pp.2807-2810, 2009.

[6] R. Cárdenes, M. Bach, Y. Chi, I. Marras, R. Luis, M. Andersson, P. Cashman, and M. Bultelle, "Multimodal evaluation for medical image segmentation", *CAIP 2007*, pp. 229-236, 2007.

An Explicit Boundary Condition Treatment of a Diffusion – Based Glioblastoma Tumor Growth Model

Stavroula G. Giatili, Nikolaos K. Uzunoglu and Georgios S. Stamatakos

***Abstract*—A careful boundary condition handling is a *sine qua non* prerequisite for a reliable diffusion based solution to the problem of clinical tumor growth and in particular glioma progression. However, to the best of our knowledge no explicit treatment of the numerical application of boundary conditions in this context has appeared in the literature as yet. Therefore, the aim of this paper is to outline a detailed numerical handling of the boundary conditions imposed by the presence of the skull in the case of glioblastoma multiforme (GBM), a highly diffusive and invasive brain tumor.**

X. Introduction

Glioblastoma multiforme (GBM) is the most common and most aggressive type of primary brain tumor in humans involving glial cells and accounting for 52% of all parenchymal brain tumor cases and 20% of all intracranial tumors. Despite improvements in cancer treatment, its overall prognosis is still very poor. Glioma cells are found in tissues surrounding the tumor, even after total resection of the tumor parts detectable by MRI scanning. The limits between tumor and normal tissue cannot be accurately determined using current technology. The poor response to treatment is partly due to the fact that GBM is highly invasive resulting to the infiltration of cancerous cells into the surrounding healthy tissue. Consequently, it is difficult to determine the underlying tumor growth dynamics across spatiotemporal dimensions. This kind of behavior contributes to the pronounced resistance to treatment and the essentially inevitable recurrence of the tumor. Many efforts have been made in order to investigate and model both the discrete aspects of glioblastoma growth and response to treatment [1] and its diffusive progression [2,3] for large clinical tumors. The diffusion equation has been applied in order to model the diffusive aspects of free tumor growth. However, to the best of our knowledge no explicit treatment of the application of boundary conditions in this context has appeared in the literature as yet. Therefore, this paper outlines a detailed numerical handling of the boundary conditions imposed by the presence of the skull in the case of glioblastoma multiforme. The execution of a number of pertinent computational scenarios has supported the correctness of the mathematical treatment presented.

Manuscript received May 30, 2010. This work was supported in part by the European Commission under the project "ContraCancrum : Clinically Oriented Translational Cancer Multilevel Modeling" (FP7-ICT-2008-2-223979.)

S. G. Giatili (giatili@otenet.gr), N. K. Uzunoglu (nnap@otenet.gr) , and G.S. Stamatakos (corresponding author, gestam@central.ntua.gr, phone: +30-210-772-2287, fax: +30-210-772-3557) are with the *In Silico* Oncology Group, Institute of Communication and Computer Systems, National Technical University of Athens, Iroon Polytechniou 9, Zografos, GR-157 80, Greece.

XI. The Diffusion based Simulation model

According to the diffusion based approach the tumor is considered a spatiotemporal distribution of continuous cell density which follows the general diffusion law. The macroscopic formulation of diffusion, leads to a partial parabolic differential equation. A single tumor cell may constitute the initial tumor within a three-dimensional medium. Tumor growth can be expressed by the following statement [2,3]:

Rate of Change of Tumor Cell Population= Diffusion (motility) of tumor cells + Net proliferation of tumor cells – Loss of tumor cells due to treatment

Additionally, simulations should address the physical processes taking place in the vicinity of the anatomic boundaries and therefore must satisfy specific constraints. The skull acts as an adiabatic boundary for the diffusion of a brain tumor whose main bulk may or may not lie at a substantial distance from it. Zero flux boundary conditions have to be applied on the skull surface. 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. Thus the first step to be taken is to select the appropriate type of boundary conditions.

In the case of a diffusive glioma lying within the brain an eventual adoption of Dirichlet boundary conditions would lead to predefined values of *tumor cell concentration* on the brain-skull boundary. In the simplest case of pure diffusion, the latter could only be achieved through the emergence of (artifact) tumor cell sources and /or sinks that would keep the boundary tumor cell

concentration values constant. Obviously such an approach would be unacceptable. On the other hand an eventual adoption of Neumann boundary conditions would impose a predefined normal derivative of the tumor cell concentration on the brain-skull boundary [3-6]. In the special case that $\frac{\partial c}{\partial n} = 0$, where c is the tumor cell concentration, adiabatic boundary conditions are implied. These are called homogeneous Neumann boundary conditions. The latter physically correspond to no net flow of tumor cells out of or into the brain region across the brain-skull boundary. In that case the diffusion region is insulated. Therefore, Neumann boundary conditions are the most appropriate for the problem addressed.

Thus if R 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 [2,3]:

$$\left\{\begin{array}{l} \dfrac{\partial c}{\partial t} = \nabla(D\nabla c) + \rho c - G(t)c \quad in\ R \\ c(\mathbf{x},0) = f(\mathbf{x}), \quad initial\ condition \\ \mathbf{n}\cdot D\nabla c = 0 \quad on \quad \partial R, \quad boundary\ condition \end{array}\right\}$$

where *c* represents cell concentration at any spatial point **x** and time *t*, *D* denotes the diffusion coefficient and represents the active motility of tumor cells, ρ represents the net rate of tumor growth including proliferation, loss and death and **n** is the unit normal to the boundary ∂R of the brain domain *R*, and *f* is a known function. The term *G(t)* accounts for the temporal profile of treatment such as radiotherapy and/or chemotherapy and as a first approximation *G(t)=k* is constant. The domain *R* is:

$$R = \{(x,y,z) \mid a < x < b, c < y < d, e < z < f\}$$

XII. Mathematical Treatment

Little progress has been made towards developing analytical solutions to the three dimensional diffusion equation with constant diffusion coefficient D when initial and boundary conditions are complex. Therefore, the study of the diffusion equation from both the analytical and even more from the numerical standpoint is still an active field of research. The objective is to approximate the exact solution to the boundary value problem at a discrete set of spatial points and convert the continuous partial differential equation into a system of discrete algebraic equations.

The first step in the problem approximation process is to make spatial and time descretization introducing a computational grid that is applied on the anatomic region of interest. Concerning the spatial descretization, the intervals $[a,b],[c,d],[e,f]$ are divided into N, M, L equal – sized sub – intervals, respectively, where N, M, L are positive integers. For convenience, let

$$\Delta x = \Delta y = \Delta z = \frac{b-a}{N} = \frac{d-c}{M} = \frac{f-e}{L} = h$$

The descretization is completed by the division of the time axis into uniform steps of length Δt and $t_n = n\Delta t$.

Having defined the computational grid, finite - difference methods are applied due to the complexity of the biological system. There are many different numerical approaches, for solving multidimensional parabolic initial – boundary value problems. The Crank-Nicolson method is considered the method of choice for many diffusion problems. It makes use of the central difference at time $t_{n+1/2}$ and is the most accurate scheme for small time steps and second - order accurate in both time and space and unconditionally stable. For the three dimensional problem the Crank - Nicolson scheme takes the form:

$$\begin{aligned} \frac{c_{i,j,k}^{t+1} - c_{i,j,k}^{t}}{\Delta t} = {} & \frac{1}{2} D\left(\frac{c_{i+1,j,k}^{t+1} - 2c_{i,j,k}^{t+1} + c_{i-1,j,k}^{t+1}}{\Delta x^2} + \frac{c_{i+1,j,k}^{t} - 2c_{i,j,k}^{t} + c_{i-1,j,k}^{t}}{\Delta x^2}\right) + \\ & \frac{1}{2} D\left(\frac{c_{i,j+1,k}^{t+1} - 2c_{i,j,k}^{t+1} + c_{i,j-1,k}^{t+1}}{\Delta y^2} + \frac{c_{i,j+1,k}^{t} - 2c_{i,j,k}^{t} + c_{i,j-1,k}^{t}}{\Delta y^2}\right) + \\ & \frac{1}{2} D\left(\frac{c_{i,j,k+1}^{t+1} - 2c_{i,j,k}^{t+1} + c_{i,j,k-1}^{t+1}}{\Delta z^2} + \frac{c_{i,j,k+1}^{t} - 2c_{i,j,k}^{t} + c_{i,j,k-1}^{t}}{\Delta z^2}\right) + \\ & \frac{1}{2}\left(\rho^{t+1} c_{i,j,k}^{t+1} - G^{t+1}(t) c_{i,j,k}^{t+1} + \rho^{t} c_{i,j,k}^{t} - G^{t}(t) c_{i,j,k}^{t}\right) \end{aligned}$$

where $c_{i,j,k}^{t}$ is the finite difference approximation of c at the grid point *(i,j,k)* at time *t*. The resulting system of equations may be written equivalently in the form $\vec{A}\vec{x} = \vec{b}$, where *x* arrow denotes a vector that contains an approximation of the solution *c* at the mesh nodes at time $t = t_n$, where $t_n = n.\Delta t$, $\Delta t > 0$.

The algorithm selected for the solution of this large sparse system is the non - stationary iterative method Conjugated Gradient method (CG) which generates a sequence of approximations that converges rapidly to the desired solution [Fig.1].

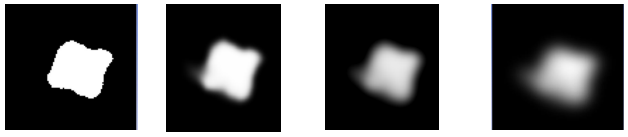


Fig. 1 A two dimensional section of a *fictitious* growing virtual glioblastoma tumor starting with well defined initial boundaries. The latter become blurred with time due to diffusion. The initial (generally arbitrary) tumor

shape is based on the boundary of the T1 gadolinium enhanced MRI imageable component of a real tumor. Several snapshots of the simulation aiming *only* at the code validation are depicted.

The boundary conditions applied refer to an arbitrarily shaped boundary. For each boundary mesh node (which lies at the center of the multi-colored structure and therefore cannot be shown) all its 6 adjacent nodes (lying towards the x+, x-, y+, y-, z+, z- directions) are considered [Fig.2].

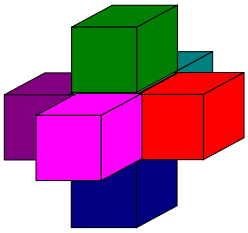


Fig 2: Boundary mesh node (non visible here) with its 6 adjacent nodes, each one lying at the center of the corresponding cubicle.

At the grid point (i,j,k) where x- belongs to the boundary:

$$-\frac{\partial c}{\partial x}\Big|_{(i,j,k)} = 0 \Rightarrow c_{i+1,j,k} = c^{x-}_{f_{i-1,j,k}}$$

At the grid point (i,j,k) where x+ belongs to the boundary:

$$\frac{\partial c}{\partial x}\Big|_{(i,j,k)} = 0 \Rightarrow c_{i-1,j,k} = c^{x+}_{f_{i+1,j,k}}$$

At the grid point (i,j,k) where y- belongs to the boundary:

$$-\frac{\partial c}{\partial y}\Big|_{(i,j,k)} = 0 \Rightarrow c_{i,j+1,k} = c^{y-}_{f_{i,j-1,k}}$$

At the grid point (i,j,k) where y+ belongs to the boundary:

$$\frac{\partial c}{\partial y}\Big|_{(i,j,k)} = 0 \Rightarrow c_{i,j-1,k} = c^{y+}_{f_{i,j+1,k}}$$

At the grid point (i,j,k) where z- belongs to the boundary:

$$-\frac{\partial c}{\partial z}\Big|_{(i,j,k)} = 0 \Rightarrow c_{i,j,k+1} = c^{z-}_{f_{i,j,k-1}}$$

At the grid point (i,j,k) where z+ belongs to the boundary:

$$\frac{\partial c}{\partial z}\Big|_{(i,j,k)} = 0 \Rightarrow c_{i,j,k-1} = c^{z+}_{f_{i,j,k+1}}$$

26 different cases of nodes having boundary node(s) as their neighbor(s) have been considered. This has led to the formulation of 26 algebraic equations (denoted by eq. (1) to (26)).

An appropriate equation out of the set of equations (1) to (26) is used for any index triplet (i,j,k) belonging to the boundary. By fixing indices i,j,k to specific values, equations (1) to (26) can produce all elementary boundary arrangements encountered in the case of an arbitrarily shaped boundary.

The execution of a number of pertinent computational scenarios using virtual tumors has supported the correctness of the mathematical treatment presented. Validation criteria include *inter alia* the prohibition of passage of tumor cells through the skull and the conservation of the total number of tumor cells within the brain in the special *fictitious* case of pure diffusion with neither cell generation nor cell death.

XIII. Conclusions

An explicit numerical treatment of the boundary conditions to be used in conjunction with a diffusion based glioblastoma tumor growth model has been presented. Through numerical experimentation simulated tumors have shown to satisfy the expected macroscopic behavior of glioblastoma multiforme including the adiabatic behavior of the skull. Such a detailed treatment of the boundary conditions may considerably contribute to the accuracy of the overall diffusion solution especially for glioblastomas having their main bulk close to the skull.

References

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

[2] K. R. Swanson, Carly Bridge, J.D. Murray, Ellsworth C. Alvord Jr. Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. Journal of the Neurological Sciences 2003; 216:1-10. K. R. Swanson, Carly Bridge, J.D. Murray, Ellsworth C. Alvord Jr. Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. Journal of the Neurological Sciences 2003; 216:1-10.

[3] J.D. Murray. Mathematical Biology II: Spatial Models and Biomedical Applications. Third Edition, Springer-Verlag, 2002; Chapter 11

[4] Medhi Dehgan. Numerical Solution of the three-dimensional advection-diffusion equation, Applied Mathematics and Computation 150 (2004) 5-19.

[5] Allan Jeffrey, Applied Partial Differential Equations-An introduction. Academic Press, An imprint of Elsevier Science, Copyright 2003; Chapter 1.

[6] Brian Bradie. A friendly Introduction to Numerical Analysis. Pearson International Edition 2006; Chapter 8,9.

VII. Computing and Multiparametric Exploration Technologies

Using the GPU for Simulating Spatiotemporal Tumor Growth

B. Liu, B; G J Clapworthy; E. Kolokotroni, G. Stamatakos

***Abstract* — Simulators of tumor growth can estimate the evolution of tumor volume and the quantity of cells as functions of time. However, the execution time of each simulation often takes several dozens of minutes (depending on the dataset resolution), which clearly prevents easy interaction. The modern graphics processing unit (GPU) is not only a powerful graphics engine but also a highly parallel programmable processor featuring peak arithmetic and memory bandwidth that substantially outpaces its CPU counterpart. In this work, by programming an NVIDIA GPU device with CUDA, we have designed algorithms to parallelise the time-consuming process of tumor simulation on the GPU, which allows the local application of basic biological rules and subsequently leads to the spatiotemporal simulation of the tumor system.**

I. INTRODUCTION

Statistical approaches to tumor growth using a Monte Carlo method have been in use for some time [1,2]. They may require thousands of times of simulations to provide a picture of the likely tumor development. Depending upon the outcome, the clinician may subsequently need to slightly tune a few parameters to obtain a new simulation result, so the interactive response time of the simulator is critical.

Ideally, for an efficient tool, each execution should be performed in seconds. However, each simulation instance will generally take several dozens of minutes, so the process cannot take place interactively, which severely limits the level of experimentation that can take place.

In the simulation method used in the *Contra Cancrum* project, a cubic discretising mesh is superimposed upon the anatomic region of interest and this is scanned every hour in order to allow the local application of basic biological rules that subsequently lead to the spatiotemporal simulation of the evolution of the tumor system. Tumor simulation is computed on these independent Geometric Cells (GCs), which provides a structure that is amenable to parallelisation. Although the intrinsically parallel architecture of the graphics processing unit (GPU) should make its use attractive in this area, no previous attempts have been made to do so – this paper describes how an existing method for spatiotemporal tumor growth was adapted for use on a GPU.

Manuscript received April 20, 2010. This work was supported in part by the European Commission within the Contra Cancrum project (FP7-IST-223979).

B. Liu and G J Clapworthy are with the University of Bedfordshire, Luton, UK (e-mails: Baoquan.Liu@beds.ac.uk, Gordon.Clapworthy@beds.ac.uk).

E. Kolokotroni and G. S. Stamatakos are with the National Technical University of Athens, Institute of Communication and Computer Systems, In Silico Oncology Group, Iroon Polytechniou 9, Zografos, GR-157 80, Greece (e-mails: ekolok@iccs.gr, gestam@central.ntua.gr).

II. THE GPU

A modern GPU is not only a powerful graphics engine but also a highly parallel programmable processor featuring peak arithmetic and memory bandwidth that substantially outpaces its CPU counterpart. A state-of-the-art GPU can perform 1.35 trillion arithmetic operations per second [3], and this represents a tremendous computational resource that can now be utilised for general purpose computing as a result of recent advances in GPU hardware and software architecture. Despite this, the GPU remains very little used in the context of the Virtual Physiological Human (VPH) – the long-term initiative funded by the European Commission of which *Contra Cancrum* forms part.

A recent survey found that, apart from *Contra Cancrum*, only 3 VPH projects made any use of GPUs and all of these used them purely for tasks related to visualisation and imaging: collision detection, image processing, etc. As a whole, VPH is ignoring the immense possibilities afforded by GPGPU (general purpose GPU programming).

CUDA (Compute Unified Device Architecture) [3] is a general-purpose parallel-computing architecture of modern NVIDIA GPUs. The CUDA programming model is based upon the concept of a kernel. A kernel is a function that is executed multiple times in parallel, each instance running in a separate thread. The threads are organised into one-, two- or three dimensional blocks, which in turn are organised into one- or two-dimensional grids. The blocks are totally independent of each other and can be executed in any order. However, threads within a block are guaranteed to be run on a single multiprocessor. This makes it possible for them to synchronise and share information efficiently using the on-chip memory, as CUDA allows all threads in a block to share data via fast on-chip shared memory to avoid redundant memory access.

III. THE APPLICATION

Using CUDA, we have designed algorithms to parallelise the time-consuming process of tumor simulation based on the biological mathematical model, in

order to accelerate this kind of numerical calculation on stream processors.

The CUDA architecture is built around a scalable array of multithreaded streaming multiprocessors. NVIDIA's current GPU consists of an array of multiprocessors, each capable of supporting up to 1024 co-resident concurrent threads. Our work decomposition is based on fine-grain task parallelism that achieves load balancing among the GPU multiprocessors.

Each GC of the discretising mesh constituting the region of interest contains a number of (biological) cells. All cells in all GCs follow the cytokinetic diagrams, which are general cytokinetic models that can be adapted for specific tumor data and drugs under consideration by adequately adjusting the corresponding simulation parameters (e.g. the probabilities of the various transitions between phases, the cell cycle durations, etc.) [1,2]. The discrete character of the simulation model enables the consideration of various exploratory initial percentages of the cells in the various equivalence classes.

The tumors are assumed to be spatially uniform and all GCs are initialised in the same way. We use parallel threads to compute for each GC independently and execute hourly scanning steps via parallel GPU kernels, which allow the local application of basic biological rules. In this way, the tumor evolution is computed by iteratively scanning the biological parameters in the GCs, and at the same time applying basic biological rules to them to spatiotemporally simulate the evolution of the tumor system. By buffering neighbouring GCs's parameters in the shared memory, we can make a massive reduction in the number of memory transactions.

At each iterative scan, we assign one thread to perform the computation associated with a single GC. To simulate tumor expansion and shrinkage, we execute the following 3 scanning steps via parallel GPU kernels:

1) *cell growth*: each thread computes all the cell-cycle phase transitions and cell deaths due to chemotherapy or radiotherapy that have occurred inside the GC.

2) *cell transfers* are computed in parallel by unloading excess cells from the current GC into the 26 adjacent GCs, using 27 subpasses that interleave the 27 neighbouring GCs at each subpass. This novel implementation takes full advantage of all GPU resources made available under the CUDA programming model. All the cell transfers among the 27 local neighbouring GCs (a box) are evaluated in a single thread, with each thread block responsible for a row of boxes. All threads in a block simultaneously iterate through the neighbouring GCs in shared memory, computing the transfers on the boxes in their individual registers. Since all threads in a block access the same shared memory location, data are broadcast to all threads by the hardware, and no bank conflict penalty is incurred.

3) *differential tumor shrinkage* is dealt with by freeing the GCs containing too few cells, or creating new GCs for differential tumor expansion. This stage also restores tumor contiguity if tumor fragmentation has taken place.

The above process is repeated continuously until completion of the time interval $T_{interval}$ (in hours) being studied.

IV. Results

We performed our experiments of free tumor growth (based on cytokinetic models) on a desktop computer with an AMD 2.3 GHz CPU and an NVIDIA GTX 285 GPU. We implemented our GPU algorithm using the CUDA programming language [3].

The performance improvement of the GPU computation compared to that on the CPU is shown in Table 1, where $time_c$ and $time_g$ are the CPU and GPU computing times, respectively.

TABLE I
PERFORMANCE IMPROVEMENT OF THE GPU OVER THE CPU

resolution of GCs	$T_{interval}$ (hours)	$time_c$ *(ms)*	$time_g$ *(ms)*	*speedup factor*
32^3	500	9852	173	56
32^3	5000	97338	1563	62
32^3	10000	197645	3121	63
64^3	5000	1169838	7905	147

It can be seen that the more demanding the calculation becomes, the greater the benefit derived from using the GPU. The speedup increases both with an increase in $T_{interval}$ and an increase in the number of GCs (that is, an increase in spatial resolution). This is due to an increase in the GPU occupancy (that is, the percentage of the hardware's ability to process threads that are actively in use) which, in essence, determines how successfully the hardware is kept busy. This can also help to reduce the effect of memory latency and dependencies.

The performance gain is particularly noticeable in the 64^3 example, where the overall processing time was reduced from 20 minutes to 8 seconds, which can radically change the approaches that users employ to investigate their data.

In terms of precision, the current GPUs fully support single precision float arithmetic and conform to the floating point standard in the same way as the SIMD units in CPUs do. Numerical codes executing on these architectures today typically yield bit-identical results, and any discrepancies are within the floating point standard [4].

Unfortunately, on our GPU (NVIDIA GTX 285), the double precision arithmetic is not very mature, so we used only single-precision float arithmetic in both the CPU and the GPU implementations. We intend, in the near future, to update our implementation into double-precision arithmetic on the newer NVIDIA Fermi GPU, such as the GTX480, which has been specifically designed to offer unprecedented performance in double precision [5], a dramatic improvement over the GT200 GPU architecture

that we currently use and we anticipate even greater performance gains when we do.

V. CONCLUSION

We have proposed a GPU-based approach to tumor growth simulation that is carefully constructed to take full advantage of the particular architecture of the GPU.

We assign GPU threads to perform tumor simulation by executing hourly scanning via parallel GPU kernels under the CUDA programming model. In this way, we can simulate cell-cycle phase transitions based on biological rules, and cell death due to chemotherapy or radiotherapy that have occurred inside each GC; we can also compute cell transfer between 27 local neighbouring GCs. Further, we can compute differential tumor shrinkage by freeing the GCs containing too few cells, and differential tumor expansion by creating new GCs. As a result, we can estimate tumor volume and the quantity of cells (stem, proliferating, etc.) as a function of time from a given set of input parameters.

Use of the GPU has proved to be an effective way of reducing the computational time to the extent that interactive data investigation can be considered for large datasets, even on commodity hardware.

REFERENCES

[1] G.S.Stamatakos, V.P.Antipas, and N.K. Uzunoglu , *A spatiotemporal, patient individualized simulation model of solid tumor response to chemotherapy in vivo: the paradigm of glioblastoma multiforme treated by temozolomide*, IEEE Transactions on Biomedical Engineering, 53(8), 1467-1477, 2006

[2] D.D.Dionysiou and G.S.Stamatakos, *Applying a 4D multiscale in vivo tumor growth model to the exploration of radiotherapy scheduling: the effects of weekend treatment gaps and p53 gene status on the response of fast growing solid tumors*, Cancer Informatics, 2, 113-121, 2006

[3] NVIDIA, *NVIDIA CUDA programming guide*, January 2010

[4] A. R. Brodtkorb, C. Dyken, T. R. Hagen, J. M. Hjelmervik, O. O. Storaasli, *State-of-the-Art in Heterogeneous Computing*, Journal of Scientific Programming, to appear

[5] NVIDIA, *NVIDIA Fermi Compute Architecture Whitepaper*, January 2010

Validating the ACGT Oncosimulator with a Grid-Supported Visualisation Environment

Aran Lunzer, Robert Belleman, Paul Melis, Juliusz Pukacki, Paweł Spychała, and Georgios Stamatakos

***Abstract*—The ACGT Oncosimulator is an integrated Grid-based system, under development within a 25-partner European-Japanese project, for patient-specific simulation of the response of a tumor and its surrounding tissue to various forms of therapy. The validation of the simulation code is an activity requiring extensive human-driven visual investigation of the influence of each of the dozens of parameters to the code, initially by comparing results from simulations carried out with different parameter values. This activity requires that users be supported in specifying simulation runs based on chosen parameter-value combinations, submitting the runs for execution on the Grid, then obtaining result visualisations that help in making the necessary comparisons. We report on our development and early use of the OncoRecipeSheet, an environment designed to meet these requirements.**

I. Introduction

One activity within the EU project ACGT[1] is the development of an Oncosimulator [5,6,1] that can simulate the behavior of tumor and normal tissue under conditions of free growth and various treatment modalities, based on a top-down multiscale simulation strategy implemented by researchers at the *In Silico* Oncology Group of the National Technical University of Athens (www.in-silico-oncology.iccs.ntua.gr). In ACGT the simulator is being used to investigate the growth and treatment of nephroblastoma and breast cancer, and runs within a secure Grid infrastructure.

The simulator code is designed to process patient-specific data. Simulation starts from a three-dimensional representation of a patient's imaged tumor, and is controlled by parameters that include pharmacodynamic factors estimated from clinical data obtained for the patient, along with dosage and timing data for the treatment to be simulated. The eventual goal is that one day a simulation like this could be used to help decide the best course of treatment for each individual cancer patient.

Before the simulator can be adopted as a tool for clinical practice, the accuracy of its predictions must be tested and tuned by extensive comparison against the actual outcomes of clinical treatments of tumors [1]. Prior even to such testing, its developers must confirm that the simulator behaves consistently and predictably as each of its inputs is varied, individually and in concert, over ranges of plausible values. We are currently addressing this latter 'validation' stage, which involves the following challenges:

The human user must be in charge. This is an application that requires human-guided visualisation rather than automated optimisation. The simulator developers must confirm, for example, that each simulation follows a credible course between its start and end conditions. This calls for inspection of the time plots of all significant elements of the simulator's internal state to check for suspicious discontinuities, and likewise of the modelled 3D tumor shape to confirm that it is not deformed in unlikely ways.

A large parameter space. The versions of the simulator we are now testing have between 35 and 40 parameters. An interface offering simultaneous access to all these parameters would be unwieldy, so we had to design a way to let users choose which parameters to view and to vary. Furthermore, since these parameters in combination delineate a space of at least 10^{20} possible results, we must help users to understand which points in that result space have been evaluated so far, and which points it would be worth evaluating next.

Comparison is key. The validation calls heavily on comparison. For a start, the simulator developers must confirm that for each step in the value of a single parameter, the results change by a reasonable increment in the predicted direction. More complex comparisons involve the interactions between multiple parameters, and here again the size of the result space militates against supporting only naive cross-product enumeration of parameter combinations: a user must be able to set up just

Manuscript received June 25, 2010. This work was supported in part by the European Commission, within the project ACGT (FP6-2005-IST-026996).

A. Lunzer works at Hokkaido University, Sapporo, Japan (corresponding author; phone: +81-11-706-7250; fax: +81-11-706-7808; e-mail: aran@meme.hokudai.ac.jp).

R. Belleman and P. Melis are at the University of Amsterdam, The Netherlands (e-mail: R.G.Belleman@uva.nl, paul@science.uva.nl).

J. Pukacki and P. Spychała are at the Poznan Supercomputing and Networking Center, Poland (e-mail: pukacki, spychala@man.poznan.pl)

G. Stamatakos is at the National Technical University of Athens, Greece (e-mail: gestam@central.ntua.gr).

[1] Advancing Clinico-Genomic Trials on Cancer, European Union FP6 Integrated Project 2006--2010; www.eu-acgt.org

those combinations that are of interest, such as p_1 with q_1, p_2 with q_1, p_3 with q_4.

Some result visualisations require interaction. The results from the Oncosimulator include representations of the tumor shape in three dimensions. These can be rendered in 2D for on-screen display, but to understand the entire solid shape a user may need to interact with the view, for example by rotating it or taking slices. Given the importance of comparisons, the interaction facilities must also support equivalent manipulation of multiple cases at the same time.

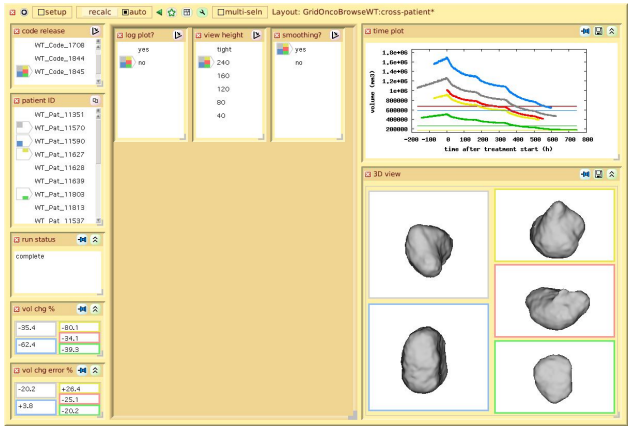


Fig.1 (Left) Viewing the default simulation results for five patients

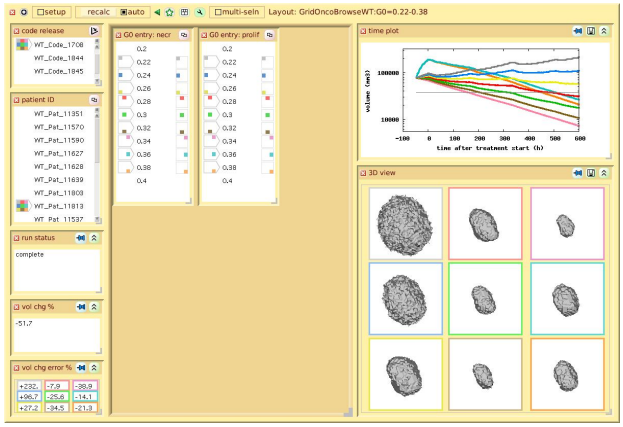


Fig.1 (Right): Results for a single patient, but nine alternative values for probability of cells entering G0.

II. Implementation

We implemented the simulator-validation interface based on the RecipeSheet [3], a spreadsheet-inspired environment that includes explicit support for users to set up and manipulate multiple alternative application states in parallel. In the case of the Oncosimulator this would allow a user to view and compare the results from simulations that use a range of alternative values for various parameters.

Like a spreadsheet, the RecipeSheet provides support for setting up custom flow-like calculations in terms of dependencies between cells. Built into the RecipeSheet are features based on the principles of subjunctive interfaces – interfaces that allow a user to explore alternatives by setting up multiple application scenarios at the same time, viewing those scenarios side by side, and manipulating them in parallel [4]. These features mean that the cells providing inputs at the start of a RecipeSheet flow (referred to as ingredients) can hold multiple values simultaneously, the user can set up alternative scenarios based on chosen combinations of those values, and the cells holding derived values will then show the results for all scenarios, colour-coded and/or spatially arranged to clarify which result arose from which scenario.

Whereas the inter-cell dependencies in a classic spreadsheet are defined using single-valued formulas, in a RecipeSheet's calculation flow the dependencies are defined using multi-input, multi-output procedures that we refer to as recipes. The main recipes used on the OncoRecipeSheet are as follows: a run-defining recipe, that accepts as inputs the parameters used to specify a simulation, and looks up the identifier of the existing run (if it indeed exists) corresponding to that specification; and visualisation recipes, that each take such a run identifier and request a visualisation, of one form or another, of the results for that run. Cases where no run has yet been performed for the specified parameter combination are handled by a further recipe, driving a run-submission tool that is external to the OncoRecipeSheet.

The Oncosimulator as a whole, being one of the demonstration applications for ACGT's secure Grid-based processing and data-sharing infrastructure, runs as a cooperation among resources distributed across many sites. The raw input data relating to each patient are maintained in pseudonymised form on a centralised data management service (DMS), and executions are requested through a Grid resource management system (GRMS); both systems are managed by Poznan Supercomputing and Networking Centre. Simulation results, once ready, are transferred from the DMS to a visualisation server at the University of Amsterdam, from which the desired visualisations can be requested. Thus the OncoRecipeSheet's visualisation recipes perform their job by communicating with the visualisation server, using a simple HTTP-level protocol, to retrieve the required results.

The snapshot on the left of Figure 1 shows a sheet on which the user has chosen to view the treatment simulations for five distinct patients' tumors. The sheet's left-hand column is dedicated to two fixed parameter cells

(for code version and patient) and three small result cells. The central region is for holding additional parameter cells, chosen by the user from two pop-up menus: one for simulation parameters, the other for parameters controlling the visualisations. All parameters not displayed as cells automatically take their default values. The result cell at top right shows a plot of tumor volume against time, with the lines from the multiple cases overlaid and colour-coded in correspondence with the markers in the patient-selection cell. At bottom right on the sheet are 3D views of the predicted final shape of the tumor in each case.

The choice and arrangement of result displays is fully configurable; the layout described here is just one possible setup of a sheet for exploring Oncosimulator results. However, although it is straightforward for a user to resize and rearrange cells on the sheet, the creation of new recipes to fetch additional aspects of the results, or to display them in alternative ways, requires specialist knowledge of the tools and protocols supported by the visualisation server. Up to now the OncoRecipeSheet developers have built all required recipes, in response to requests from the developers of the simulator code. In all, Oncosimulator-specific recipes and the mechanisms that they invoke amount to some 3,000 lines of Smalltalk code, the bulk of this being for communicating with the DMS and GRMS to manage the repository of completed simulations, to schedule new runs, and to orchestrate these requests so that users at multiple sites can work simultaneously without resource-demand clashes.

Further details of the OncoRecipeSheet are given in [2].

III. Initial Findings

We installed a preliminary version of the Onco-RecipeSheet at two sites in the summer of 2009. The sheet was already set up with all necessary recipes for a basic exploration of simulator results as presented above; the users are not expected to make changes to this setup themselves. However, users are free to change the layout of the sheet, for example if needed to suit the size of screen being used. They are also able to customise for themselves, using XML-based configuration files in the Onco-RecipeSheet's working directory, the ranges of values offered for each simulation parameter. For example, in the distributed configuration the range of cell-cycle durations is set to use increments of 2 hours; a user wanting to explore the effects of 1-hour differences could edit the XML to set up the desired range.

We have so far gathered roughly 3,000 simulation results for the two cancer types being examined. Although in general the simulator has proven to be robust and predictable across a broad range of parameter settings, the ability to run and to inspect large numbers of simulations has already helped us to uncover some unexpected behaviors. In one set of runs we found that results differed depending on which Grid node happened to be used for the execution; this led to the discovery of a previously undetected bug within the code. Combinations of runs such as those shown on the right in Figure 1 have also been informative. In this case the results are for a single patient, under a range of values for parameters that set the probabilities for certain cell-state transitions. The tumor volume plots on this sheet show a dramatic discontinuity in response for higher probability values. Although these values are known to represent a biologically impossible situation, it is valuable to observe how the simulator behaves when driven to these extremes.

Meanwhile the generation of simulation results over the past year has constituted a useful stress-test of the ACGT Grid infrastructure, bringing to light many issues that we were then able to resolve, thus improving the reliability and performance of the overall system.

IV. The Way Forward

As emphasised in the Introduction, these are still very early days in a process towards development of an Oncosimulator that might one day be used in clinical practice. We believe that being able to obtain and compare results from multiple alternative simulations, using facilities like those of the OncoRecipeSheet, will continue to be crucial in future stages of this process. For example, the results from the existing simulator show that apparent increases or decreases in therapy effectiveness can be induced by variation in a number of the model's parameters, independently or in concert. Being able to run multiple simulations will help in exploring the diverse parameter-value combinations that would all lead to similar predictions, and thus in setting confidence levels regarding our ability to match predictions against observed clinical results. Further in the future, an interface that shows a selection of the diverse treatment outcomes that would be predicted in the face of typical (and unavoidable) inaccuracies in patient-data measurements is likely to be more credible than a 'black box' that produces a single result. Such a tool would also help in educating clinicians about what these simulations can tell us... and what they can't.

Acknowledgements

We are grateful for the support of all our colleagues in our home institutes and across ACGT, but in particular we thank Dimitra Dionysiou, Eleni Georgiadi and Eleni Kolokotroni of the In Silico Oncology Group at the

National Technical University of Athens, Norbert Graf and Alexander Hoppe at the University Hospital of the Saarland, Marius Erdt at Fraunhofer IGD, and Julien Jacques at IRISA-INRIA.

The first author is supported by Hokkaido University's Global Centre of Excellence (GCOE) programme for Next-Generation Information Technology Based on Knowledge Discovery and Knowledge Federation.

REFERENCES

[1] N. Graf, A. Hoppe, E. Georgiadi, R. Belleman, C. Desmedt, D. Dionysiou, M. Erdt, J. Jacques, E. Kolokotroni, A. Lunzer, M. Tsiknakis, G. Stamatakos (2009) In Silico Oncology for Clinical Decision Making in the Context of Nephroblastoma. *Klinische Pediatrie* 221:141–149.

[2] A. Lunzer, R. Belleman, P. Melis, G. Stamatakos (2010) Preparing, Exploring and Comparing Cancer Simulation Results Within a Large Parameter Space. In *Proceedings of 3rd International Symposium on Information Visualization in Biomedical Informatics (IVBI)*, London, UK, Jul 2010.

[3] A. Lunzer, K. Hornbæk. RecipeSheet: Creating, Combining and Controlling Information Processors. In *Proceedings of the 19th Annual ACM Symposium on User interface Software and Technology (UIST '06)*, Montreux, Switzerland, Oct 2006, 145–153

[4] A. Lunzer, K. Hornbæk. Subjunctive interfaces: Extending applications to support parallel setup, viewing and control of alternative scenarios. *ACM Trans. Comput.-Human Interact.* 14, 4, Art. 17 (Jan 2008), 1–44.

[5] G.S. Stamatakos, V.P. Antipas, N.K. Uzunoglu (2006) A spatiotemporal, patient individualized simulation model of solid tumor response to chemotherapy in vivo: the paradigm of glioblastoma multiforme treated by temozolomide. *IEEE Trans Biomed Eng*. 53(8), 1467–77, Aug 2006.

[6] G.S. Stamatakos, D.D. Dionysiou, N.M. Graf, N.A. Sofra, C. Desmedt, A. Hoppe, N. Uzunoglu, M. Tsiknakis (2007) . 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, Proceedings of the 29th Annual International Conference of the IEEE EMBS*, Aug 23-26, SuB07.1:6628–6631, Lyon, France, 2007.

VIII. Image Processing for Tumor Modeling

G. Stamatakos and D. Dionysiou (Eds): *Proc. 4th Int. Adv. Res. Workshop on In Silico Oncology and Cancer Investigation (4th IARWISOCI) – The ContraCancrum Workshop*, Athens, Greece, Sept. 8-9, 2010 (www.4th-iarwisoci.iccs.ntua.gr)

Tumor segmentation: The impact of standardized signal intensity histograms in glioblastoma

Jonathan Zepp, Norbert Graf, Emmanouil Skounakis, Rainer Bohle, Eckart Meese, Holger Stenzhorn, Yoo-Jin Kim, Christina Farmaki, Vangelis Sakkalis, Wolfgang Reith, Georgios Stamatakos, Konstantinos Marias

***Abstract* — In silico oncology is anticipated to gain a more individualized treatment for patients with cancer. To run in silico oncology models data from individual patients are essential. The more and the more accurate these data are the more precise the results of the in silico oncology models will be. Imaging studies are used to calculate tumor volume and define vital, necrotic and cystic areas within a tumor. Though the visual interpretation of magnetic resonance (MR) images is based on qualitative observation of variation in signal intensity a correlation of signal intensities with histological features of a tumor is not possible. Quantitative methods are needed for reliable follow-up or inter-individual studies. Using DoctorEye tumors can be easily rendered and histograms of the signal intensities within a tumor as well as mean and median signal intensities are calculated. In gliomas the histogram of signal intensities of cerebrospinal fluid is used as a reference for standardization of signal intensities.**

Our results in gliomas suggest that these histograms add value for a better description of tumors for the use in insilico oncology models.

I. INTRODUCTION

THE goal in cancer research is to find more individualized treatments leading to higher cure rates. In silico oncology is one of several approaches to aim this goal. MR images are crucial in finding and describing tumors. Their interpretation in respect to response is much more difficult. Not only because there is no correlation between signal intensities and histological features available there is also no clear demarcation between tumor and healthy tissue at least in gliomas possible to define. These restrictions hamper inter- and intra-individual studies and lead to imprecise data used in *in silico oncology* models.

N. Graf: University of the Saarland, Department of Pediatric Oncology and Hematology, D-66421 Homburg, Germany (Tel.: 0049 6841 1628397, Fax: 0049 6841 1628302, Email: graf@uks.eu
J. Zepp, H. Stenzhorn: University of the Saarland, Department of Pediatric Oncology and Hematology, D-66421 Homburg, Germany
R. Bohle, Y-J. Kim: University of the Saarland, Department of Pathology, D-66421 Homburg, Germany
E. Meese: University of the Saarland, Department of Human Genetics, D-66421 Homburg, Germany
E. Skounakis, C. Farmaki, V. Sakkalis, K. Marias: Institute of Computer Science at FORTH, Vassilika Vouton, GR-70013 Heraklion, Crete, Greece
W. Reith: University of the Saarland, Department of Neuroradiology, D-66421 Homburg, Germany
G. Stamatakos: the National Technical University of Athens, Institute of Communications and Computer Systems, In Silico Oncology Group, Iroon Polytechniou 9, Zografos, Greece

This work is supported by EC ContraCancrum Project: ICT-2007-223979

II. METHOD

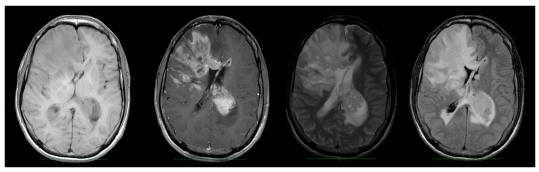
MR Images from 20 patients with glioblastoma were analyzed during their individual follow-up including T1, T1 with gadolinium contrast and T2 modalities. Signal intensities of the cerebrospinal fluid were used as reference values[1] for standardization of signal intensities within the tumor. Tumor volumes were calculated after rendering of the tumor using DoctorEye as an open source tool under the GNU General Public License [2,3]. DoctorEye provides the calculation of histograms of segmented areas and the corresponding mean and median values of signal intensities. A comparison of the shape of the histograms, the mean and median values for each modality were done in individual cases during their follow-up and between different patients. The software program SPSS was used for statistical analysis.

III. RESULTS

It is well known that different MR modalities show different shapes of a single glioblastoma making correct segmentations of tumors nearly impossible (fig.1). To overcome this problem two different segmentation methods were investigated in 20 patients with glioblastoma using histograms of the selected areas for describing the tumor. The first method included the rendering of the tumor as precisely as possible whereas the second method included the whole tumor in a square with surrounding normal tissues (fig. 2).

It can be shown that histograms of signal intensities between cerebrospinal fluid (CSF), vital tumor, necrotic and cystic areas within the tumor vary significantly in all modalities analyzed (fig. 3). Using combinations of histograms from different modalities the tumor can be described in a much better way than by calculating solely the tumor volume. Even the second method without using precise tumor segmentations can describe changes within the tumor during the follow-up of single patients that are

correlated to treatment response or progression. The higher the standardized median and mean values of signal intensities in T1 are getting during the follow-up in a single patient the more likely the patient suffers from progression of disease. If these values are going down the more likely a tumor response can be diagnosed.



T1 T1 + contrast T2 T2 flair

Figure 1: Different modalities of MR images in a single patient with glioblastoma at the same time

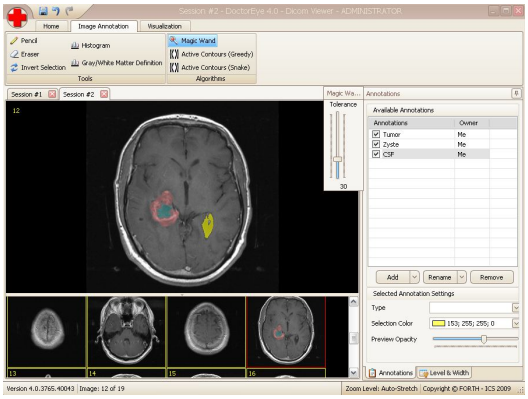


Figure 2: Segmentation of a glioblastoma using DoctorEye. Method 1 shows the vital tumor in red and the necrotic area in green For Method 2 the blue square is used. Cerebrospinal fluid is shown in yellow.

Figure 3: Histograms of 5 different patients with glioblastoma (T1 modality) showing vital and necrotic tumor areas as well as the histogram of the corresponding CSF for standardization purposes.

IV. Discussion

Application of the MR technology in medical issues is a few decades old, but the prediction of tumor response in glioblastoma using simulation models based on MR technologies is investigated by few groups only[4.5.6]. A publication done by Chen et al in 2010[7] shows, that such simulations are able to successfully predict the region of recurrence in glioblastoma. Precise data from imaging studies are of utmost importance to gain such results in *in silico oncology* models. The better these data are the more accurate results can be predicted[7,8]. For the validation of the models segmentation of the tumor during follow-up is needed. A correlation between tumor texture and signal intensities in MRI expressed by histograms of signal intensities is a step forward in the right direction. This method can be further developed by taking the localization of the different signal intensities into account, which is not the case yet. Nevertheless the use of such data in *in silico oncology* models will be tested and their impact analyzed in future top-down models of glioblastoma.

V. REFERENCES

1. Luoma K, Raininko R, Nummi P, Luukkonen R: Is the signal intensity of cerebrospinal fluid constant? Intensity measurements with high and low field magnetic resonance imagers. Magnetic Resonance Imaging 11:549-555, 1993
2. DoctorEye: http://biomodeling.ics.forth.gr.
3. 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
4. Wang Z, Deisboeck T: Computational modeling of brain tumors: discrete continuum or hxbrid? Sci Model Simul 15:381-393, 2008
5. Zhang, Wang Z, Sagotsky J Deisboek T: Multiscale agent-based cancer modeling. J MAth Biol 58:545-559, 2009
6. Chen LL, Ulmer S, Deisboeck TS: An agent-based model identifies MRI regions of probable tumor invasion in a patient with glioblastoma. Phys Med Biol 55:329-338, 2010
7. 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.
8. Stamatakos GS, Antipas VP, Uzunoglu NK: A spatiotemporal, patient individualized simulation model of solid tumor response to chemotherapy in vivo: the paradigm of glioblastoma multiforme treated by temozolomide. IEEE Trans Biomed Eng. 2006 Aug;53(8):1467-77.

Image Processing for *in-silico* Oncology and Lung Cancer

Astrid Franz, Roland Opfer, Steffen Renisch, Rüdiger Grewer, and Jörg Sabczynski

***Abstract*—We present basic image processing tools needed as input for lung tumor simulations and for tumor model validation. Assuming a number of PET/CT scans from the same patient are given, our lesion tracking algorithm registers all PET/CT pairs and all pairs of two consecutive CT scans. Requiring little user interaction by a few mouse clicks for initializing, the tumor is automatically segmented in the base-line CT and PET images. These segmentation results are propagated to all follow-up scans, serving as initialization for the subsequent automatic segmentation. The segmentation results in both modalities can be combined to classify the tumor tissue, needed in the succeeding tumor growth simulation.**

I. Introduction

Cancer is a complex disease involving processes at several levels of biocomplexity. Biomathematical models can on the one hand help to understand basic phenomena and on the other hand help to develop an individualized therapy taking into account the available information on an individual patient's disease. The ultimate goal of *in-silico* oncology is thus providing a clinical decision support tool by simulating *in-vivo* tumors and predicting the effect of potential therapies on the outcome.

While the simulation of *in-vitro* tumors and tumorlets is a necessary first step in order to create an understanding of the behavior of the underlying models, concepts for adapting *in-silico* models to an individual patient must be developed.

Many conceptual difficulties arise when dealing with real patient data. Apart from ethical considerations, the large variability of tumors on the microscopic level (i.e. genetic level) as well as on the macroscopic level (morphological level) must be taken into account by any simulation and prevents a straight forward implementation of *in-silico* models.

Image analysis of medical images is thus an unavoidable step in the simulation chain if clinical data is used. Image analysis methods provide starting values as well as boundary conditions for *in-silico* simulations by extracting spatial information on the shape and location of the tumor and the surrounding healthy tissue from the available image data. Furthermore, image analysis methods allow for model validation by quantitatively comparing model predictions with the actual tumor development and the actual normal tissue behavior following the selected treatment. The main image analysis tasks for *in-silico* oncology are thus image segmentation and image registration.

Image segmentation is a method to identify areas of equivalent tissue with respect to an underlying property. With the help of image segmentation methods, regions belonging to an organ, to a tumor, and tumor regions with similar tracer uptake, etc. can be identified. These regions are then used as the input for the simulation proper.

However, the necessary information for initializing realistic tumor models cannot be provided by a single imaging modality alone. In the example application of lung cancer modeling described in this paper, we deal therefore with CT images showing the lung anatomy as well as the tumor morphology and with FDG PET images showing the metabolic activity of the tumor. These images from different imaging modalities with different contrast must be aligned (registered), before they can be used as input for the simulation. Furthermore, for the purpose of validation, i.e. comparison of the *in-silico* model prediction with the actual *in-vivo* tumor behavior, medical images from different time points (e.g. before, during, and after therapy) must be registered to allow for a quantitative assessment of the change of functional and morphological parameters.

Manually aligning multimodal medical images and segmenting tumor regions as well as patient anatomy is a tedious task. State of the art image processing and image analysis methods help the biomathematical researcher in developing and validating *in-silico* methods in oncology. Unfortunately, a complete fully automatic suite of image analysis tools is error prone and thus not feasible today. In this paper we thus present for the example application of lung cancer methods for image segmentation and registration, which require only a minimal amount of user interaction [1].

Manuscript received May 4, 2010. This work was supported by the European Union under Grant FP7--IST-223979 (Clinically Oriented Translational Cancer Multilevel Modeling).

All authors are with Philips Research Europe – Hamburg, Röntgenstraße 24-26, 22335 Hamburg, Germany (e-mails: astrid.franz@philips.com, roland.opfer@philips.com, steffen.renisch@philips.com, ruediger.grewer@philips.com, joerg.sabczynski@philips.com), corresponding author: Jörg Sabczynski (phone: +49-40-5078-2047; fax: +49-40-5078-2510; e-mail: joerg.sabczynski@philips.com).

II. Data fusion and registration of lung images

For investigating lung cancer, we are given a number of PET/CT scans from the same patient. These images contain not only the tumor, but the whole thorax or even the whole body. The alignment of these images is required to be accurate in a region around the tumor only. Therefore we decided to use a combination of a pre-computed fast global affine registration with an on-demand local block matching to identify corresponding lesions [2]. The user marks a lesion in the base-line CT image (or in the overlaid view of the base-line PET and CT scans). The algorithm then finds the corresponding lesions in the follow-up scans automatically. Our proposed lesion tracking algorithm consists of three basic steps.

A. Global rigid registration

We apply a fast multi-resolution affine registration algorithm [3]. As similarity measure we use cross correlation between the intensity values. We compute registrations between all PET/CT pairs and between all pairs of two consecutive CT scans. Registrations between two arbitrary images from that series are defined by chaining up these global registrations.

B. Block matching

Whenever the user clicks on a lesion in the base-line CT image, a volume of interest (VOI) is cropped around this point. The VOI is propagated to the follow-up CT scans according to the global registration. The center of the VOI is then varied in the follow-up CT scans, while optimizing the cross-correlation within the VOI.

C. Local lesion search in the PET scans

In some cases the block matching can fail. The morphology near the lesion in the CT scans can change during the course of the therapy. In these cases block matching may not find the correct position of the corresponding lesions. To improve the result, we combine Steps *A* and *B* with local lesion search in the PET scans [2].

III. Segmentation of lung images

A. Segmentation of CT images

The task to segment lung tumors in CT images in three dimensions is not trivial. There are many papers and methods on lung nodule segmentation [3,4,5], but they can not directly be applied to lung tumors [1].

In [6] we have suggested to use a model-based segmentation approach [7,8]. However, the high variability in shape and size of lung tumors makes it difficult to define a general tumor surface model. Moreover, partial connection with the lung wall often occurs with little or no contrast between tumor tissue and outer-lung regions. Therefore we decided to apply a semi-automatic algorithm to segment the lung tumors, consisting of two steps:

1) Interactive definition of initial tumor surface mesh

We have developed a new technique for a fast, flexible, and intuitive 3D definition of initial tumor surface meshes. Points on the boundary of the tumor are marked by mouse clicks. After each mouse click we compute a sphere which best approximates the user-defined points. All points are projected onto this sphere, and the distances to the sphere are computed. These distances are then interpolated using radial basis functions, resulting in an interpolating deformed sphere through the user-defined points.

2) Model adaptation

To improve the accuracy of the result from the previous step we apply the well known model based segmentation methods. We convert the deformed sphere from the previous step into a triangular mesh. The adaption is an iterative process which optimizes the influence of shape constraints, given by the triangular mesh, and features of the grayscale image in each step.

B. Segmentation of PET images

The FDG PET based segmentation provides the parts of the body with a high metabolic activity. In order to segment the ”hot areas” from the PET image we have used standardized uptake value (SUV) prioritized region growing. The user can choose either an absolute SUV threshold or a percentage of the maximum SUV in the lesion as threshold. From the segmented hot areas in the PET images we can compute the maximum SUV, mean SUV, the volume of the lesion, and the total glycolytic volume (TGV) [9,10,11].

C. Combining information of both modalities

We propose to combine a CT based segmentation with a functional segmentation from the PET images. In this way we obtain a classification of the tumor into different tissue types. The CT based segmentation provides all parts of the lung belonging to the tumor whereas the PET based segmentation provides the parts with a high metabolic activity. Combining both segmentation results can yield a classification of the tumor tissue, as active and necrotic areas. This is a very important information for the subsequent tumor growth modeling.

References

[1] R. Opfer, et al, “ContraCancrum: Segmentation of different tissue types in lung tumors from PET/CT follow-up scans”, *IFMBE Proceedings, World Congress on Medical Physics and Biomedical Engineering*, 2009:25:1920-1923.

[2] R. Opfer et al, “Automatic lesion tracking for a PET/CT based computer aided cancer therapy monitoring system”, *Proceedings of SPIE Medical Imaging*, 2008:691513.

[3] J. M. Kuhnigk et al, “Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans”, *IEEE Transactions on Medical Imaging*, 2006:25:417-434.

[4] R. Opfer and R. Wiemker, “A new general tumor segmentation framework based on radial basis function energy minimization with a

validation study on LIDC lung nodules", *Proceedings of SPIE Medical Imaging*, 2007:651217.
[5] R. Wiemker et al, "Aspects of computer-aided detection (CAD) and volumetry of pulmonary nodules using multislice CT", *British Journal of Radiology*,. 2005:78:S46.
[6] R. Opfer et al, "Follow-up segmentation of lung tumors in PET and CT data", *Proceedings of SPIE Medical Imaging*, 2009:72600X.
[7] V. Pekar et al, "Shape–model–based adaptation of 3D deformable meshes for segmentation of medical images", *Proceedings of SPIE Medical Imaging*, 2001:4322:281–289.
[8] M. R. Kaus, J. von Berg, J. Weese, W. Niessen, and V. Pekar, "Automated segmentation of the left ventricle in cardiac MRI", *Proceedings of MICCAI*. 2003:8:245–254.
[9] R. J. Francis et al, "Early prediction of response to chemotherapy and survival in malignant pleural mesothelioma using a novel semiautomated 3-dimensional volume-based analysis of serial 18F-FDG PET scans", *Journal of Nuclear Medicine*, 2007:48:1449.
[10] J. A. Boucek et al, "Assessment of tumor response with 18F-fluorodeoxyglucose positron emission tomography using three-dimensional measures compared to SUV max - a phantom study". *Physics in Medicine and Biology*. 2008:53:4213–4230.
[11] M. Soret, S. L. Bacharach, and I. Buvat, "Partial-volume effect in PET tumor imaging", *Journal of Nuclear Medicine*, 2007:48:932-945.

IX. Multiscale Data and Simulation Tool Handling Technologies

G. Stamatakos and D. Dionysiou (Eds): *Proc. 4th Int. Adv. Res. Workshop on In Silico Oncology and Cancer Investigation (4th IARWISOCI) – The ContraCancrum Workshop*, Athens, Greece, Sept. 8-9, 2010 (www.4th-iarwisoci.iccs.ntua.gr)

A framework supporting sharing and reuse of data and tools in translational cancer research: Lessons learned for VPH research

M.Tsiknakis, S. Sfakianakis, G. Zacharioudakis, L. Koumakis

***Abstract*— The Virtual Physiological Human is an ambitious initiative that aims at studying the human body as a complex integrated system. The challenges introduced by such an endeavour are multiple and diverse. In this short paper we aim to highlight these challenges and also to present technologies and methodologies that when employed can greatly facilitate the integration and sharing of the data and the analytical tools.**

I. Introduction

THE main objective of the Virtual Physiological Human (VPH) initiative is to develop a systemic approach that avoids the pitfalls of subdividing biological systems in any particular way (e.g. by anatomy or discipline), but instead to follow a holistic approach by investigating the human body as an integrated (though hugely complex) system [1].

The challenges introduced by the VPH research objectives are staggering both in terms of methodological approaches and the IT infrastructure (middleware and integration platforms, tools, model and data repositories, standardized descriptions, etc.) that will support the research activities. The VPH Toolkit **Error! Reference source not found.**, a key deliverable of the VPH Network of Excellence (NoE) project, is a computational and data management framework, currently under development, which aims to provide the needed infrastructure ranging from standardized model markup languages and data bases to APIs, libraries, and tools for accessing, processing, and viewing data and models.

This is a herculean task when considered in the full spectrum of the requirements, the complexity of the VPH idea in general, the horizontal and vertical integration of data, the algorithmic and computational complexity of the models describing the biophysiological processes and the tools implementing these models or processing/mining their results, etc. Nevertheless it is the aim of this paper to argue that there are state-of-art technological solutions alleviate these problems. In particular we focus on the technological frameworks for dealing with the issues of the multilevel, distributed data integration and the semantic annotation of the analysis and knowledge extraction tools. This experience comes from the design of an ontology driven, semantic grid services infrastructure for multi-centric, post-genomic clinical trials in the context of the EU-funded project ACGT[3].

Manuscript received May, 19 2010. This work is supported in part by the European Commission in the context of the ACGT EU project (FP6/2004/IST-026996).

Authors are with the Institute of Computer Science, Foundation for Research and Technology – Hellas (ICS-FORTH), Heraklion, Crete, Greece

II. Data Integration

Sharing of data to the VPH scale, both vertically and horizontally, requires proper annotation of their semantics, quality, provenance, and meaning. Semantic data integration is necessary in our view in order to address a number of requirements:

- Data heterogeneity. The requirements for data management in the context of the VPH research include not only dealing with big volume of the data but also to harmonize diverse data models and schemas. This diversity is an essential characteristic of the VPH application domain where usually data sources ranging from the molecular, cellular, organ, and population levels.
- Data complexity. Imposition of abstraction layers facilitates the high level view of complex systems and subsequently the understanding of these systems. To this end semantics can play an important role of summarizing the important information during data curation, fusion planning, training, etc.
- Mediation and transformation. Information about the type, structure, and "meaning" of the data is valuable when different tools and services produce and consume nearly similar, at the conceptual level, data sets. In these use cases automatic transformation services can deal with superficial data inconsistencies.

- Provenance and data lineage. The accumulation of large data sets, which are either primitive, i.e. collected from the primary data sources (e.g. patients' data), or generated by some processing tools, should be complemented by the mechanisms that keep track of their lineage (history) and provenance. The provenance information can answer questions like how the data were produced, what are their origin, what were the processing steps that gave rise to their production and their parameters, what are the relationships with other data used in these processes, etc. This information is also important for the understanding and the reproducibility of scientific results and can facilitate a more objective quality evaluation.

- Data discovery and reuse. Semantics based discovery of data sets is only possible when the data have been annotated and described using detailed domain specific terms and vocabularies. The semantic descriptions of the data, by capturing their origin and purpose, can further guide their reuse assuming of course that potential security and privacy constraints are satisfied.

At the infrastructure level data integration usually requires a proper architectural design and this is an extensively surveyed topic in the database world [4]. Mediator architectures are also well studied, see for example [5]. Nevertheless in the recent years web architectures have been in the spotlight due to the emergence of the Semantic Web [6]. The Semantic Web builds upon the existing World Wide Web but instead of documents it aims to connect, in a machine-readable giant graph, "resources" i.e. both tangible and abstract "things" like data, persons, organizations, concepts, etc. This is possible by the introduction of a generic and universal data model (RDF) that abstracts from the syntactic nature of the existing web (e.g. HTML, XML) and instead focuses on the semantic properties and relationships of the represented resources. Based on the sound ground of RDF, additional abstractions and facilities are introduced: an SQL-like query language (SPARQL) for the Semantic Web, vocabularies and ontologies (RDFSchema, OWL), rules (RIF), etc.

The Semantic Web technologies make also possible to define domain ontologies using the RDF data model. Ontologies are vital for the semantic based data integration because they not only define a "schema" for the information exchanged but also allow more intelligent automated behavior like reasoning and inference. If a global domain ontology is in place then a lot of problems introduced by the data alignment, harmonization, and fusion can be eliminated or at least reduced [7]. Nevertheless in the context of the VPH research it is highly improbable that a single unique ontology will ever be able to capture all the relevant aspects and modeling levels. It is however the web based nature of the Semantic Web that makes possible the interlinking [8] between data and ontologies so that they can be reused, along side with the "open world" assumption of RDF and OWL that facilitate ontology alignment.

III. Sharing and Reuse of Tools

Tools are the other necessary ingredient of a VPH framework. If data are the "information bearers", tools are the "agents" that consume, process, and possibly produce additional data. Making available and sharing these tools are therefore quite important scenarios of use. In order to realize these scenarios semantics related metadata are again necessary and can be generally classified in:

- Data semantics, which defines the data model for the input and output data of the tools.

- Functional Descriptions define functionality and its capabilities, i.e. what a tool provides to its callers.

- Non-Functional Descriptions define additional aspects of the tool implementation and environment such as "quality of service" (performance, throughput, accuracy, etc) or policies, e.g. security.

- Behavioral Descriptions define the external and internal behavior of the tool. The externally visible behavior is for example the "protocol" the client has to follow when using the tool, e.g. the sequence of actions. On the other hand the internal behavior is related to the way the tool is implemented, e.g. is it uses additional resources like external data sources.

- Technical Descriptions define developer related details, such as message serializations and formats, communication protocols, and physical service access points.

To ease the publication and sharing of analytical tools a registry component needs to be in place as the authoritative point of reference for registering new tools and for searching and browsing the existing ones. Integration with the web opens again a lot of possibilities such as supporting user collaboration, social interactions, custom classification through "tags" etc. myExperiment [9] is one such web based environment for the sharing of scientific workflows.

IV. Conclusion

The VPH roadmap includes highly ambitious goals that require sophisticated methodologies and flexible, data and user driven infrastructures. The need for semantics based description for data and tools is evident especially in the

VPH domain where a number of disciplines meet. Nevertheless there are a number of standards, technologies, components, and tools to facilitate this endeavour:

- Semantic Web technologies provide enough of the needed infrastructure to build semantics based publication, discovery, and reuse of data and tools
- "Web2.0" or social web infrastructures foster the user empowerment by allowing them to have a more active role on how the information is shared and maintained.

REFERENCES

[1] J. Fenner, B. Brook, G. Clapworthy, P. Coveney, V. Feipel, H. Gregersen, D. Hose, P. Kohl, P. Lawford, K. McCormack *et al.*, "The EuroPhysiome, STEP and a roadmap for the virtual physiological human," *Philosophical Transactions-A-Mathematical Physical and Engineering Sciences*, vol. 366, no. 1878, pp. 2979–3000, 2008.

[2] D. Gavaghan, P. Coveney, and P. Kohl, "The virtual physiological human: tools and applications I," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 367, no. 1895, p. 1817, 2009.

[3] M. Tsiknakis, M. Brochhausen, J. Nabrzyski, J. Pucacki, S. Sfakianakis, G. Potamias, C. Desmedt, and D. Kafetzopoulos, "A semantic grid infrastructure enabling integrated access and analysis of multilevel biomedical data in support of postgenomic clinical trials on cancer," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 12, no. 2, pp. 205 –217, march 2008.

[4] A. Halevy, "Answering queries using views: A survey," *The VLDB Journal*, vol. 10, no. 4, pp. 270–294, 2001.

[5] L. Martín, E. Bonsma, A. Anguita, J. Vrijnsen, M. García-Remesal, J. Crespo, M. Tsiknakis, and V. Maojo, *Data Access and Management in ACGT: Tools to solve syntactic and semantic heterogeneities between clinical and image databases*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2007, vol. 4802, pp. 24–33.

[6] N. Shadbolt, W. Hall, and T. Berners-Lee, "The semantic web revisited," *Intelligent Systems, IEEE*, vol. 21, no. 3, pp. 96 –101, jan.-feb. 2006.

[7] M. Brochhausen, A. Spear, C. Cocos, G. Weiler, L. Martín, A. Anguita, H. Stenzhorn, E. Daskalaki, F. Schera, U. Schwarz *et al.*, "The ACGT Master Ontology and its Applications-Towards an Ontology-Driven Cancer Research and Management System," *Journal of Biomedical Informatics*, 2010.

[8] C. Bizer, T. Heath, K. Idehen, and T. Berners-Lee, "Linked data on the web (ldow2008)," in *WWW '08: Proceeding of the 17th international conference on World Wide Web*.1em plus 0.5em minus 0.4emNew York, NY, USA: ACM, 2008, pp. 1265–1266.

[9] D. De Roure, C. Goble, and R. Stevens, "The design and realisation of the *my*Experiment Virtual Research Environment for social sharing of workflows," *Future Generation Computer Systems*, vol. 25, no. 5, pp. 561–567, 2009.

X. Markup Languages and Collaborative Systems for Cancer Model Development

G. Stamatakos and D. Dionysiou (Eds): *Proc. 4th Int. Adv. Res. Workshop on In Silico Oncology and Cancer Investigation (4th IARWISOCI) – The ContraCancrum Workshop*, Athens, Greece, Sept. 8-9, 2010 (www.4th-iarwisoci.iccs.ntua.gr)

Markup Languages for *In Silico* Oncology

David Johnson, Jonathan Cooper, and Steve McKeever

***Abstract*—*In silico* models of cancer progression are numerous and diverse. Integration of different cancer models into virtual research environments and computational frameworks require the models themselves to be interoperable across the research community. In this paper we describe a vision to create a common computational framework in which these models can be formulated and stored in online repositories, such as the Center for the Development of a Virtual Tumor's (CViT) Digital Model Repository (DMR).**

I. Introduction

THE European Commission (EC) funded Transatlantic Tumor MOdel Repositories (TUMOR) project aims to develop a platform for creating a collaborative research community and enabling the clinical application of cancer models. A key component of the platform will be a cancer model repository for storing and accessing models developed by two other major EC projects, ACGT (Advancing Clinico Genomic Trials on Cancer) [1] and Contra Cancrum [2]. The 'transatlantic' component of the project is to design the European repository to interoperate with the US-based CViT DMR [3] ultimately providing an international research environment that will connect experts in the United States and Europe. Facilitating exchange and reuse of cancer models from both perspectives is key to advancing the state-of-the-art in oncology.

II. Motivation

Currently, the CViT DMR allows data files and executables of cancer models to be uploaded. However, there is no interoperability between these entries as there are no agreed standards on how such codes should be written. Cancer models are developed and implemented by hand, and require domain expertise in order to manipulate and evaluate simulation runs. In practice this means that there is no reuse of code or provision for coupling models that severely reduces the scope for collaborative developments. The heart modeling and systems biology communities have begun to address both issues through the creation of domain specific markup languages such as CellML [4] and SBML [5]. Until now however, these languages have not had the mathematical sophistication required by the cancer modeling community. With the development of FieldML [6] and the latest version of CellML, the required expressivity will soon be available.

We have begun to integrate these heart models into a database that will allow for interoperability and model reuse [7]. We aim to provide a similar level of functionality to the cancer modeling community, and for online model repositories, such as the DMR, by encoding tumor models in SMBL/CellML/FieldML. Existing models, however, will also be supported using annotations (namely our construction metadata discussed below) and wrapped so that they fit into overarching workflows. Re-coding will not be necessary and a degree of interoperability will be supported for existing codes.

The motivation for modeling markup is two-fold:

- Firstly we want to describe the implementation of these cancer models in an abstract manner that is not tied to any particular programming notation.
- Secondly we want to be able to couple our models. This kind of development has been undertaken in the context of heart modeling [8] but, to our knowledge, is not available as yet for cancer modeling. This has led to a bewildering proliferation of cancer models, most of which cannot easily be used by different research groups.

III. Markup for Cancer Models

The Oxford University Computing Laboratory has substantial experience in developing such frameworks in the context of multi-scale modeling [9]. We aim to transfer this knowledge across to cancer modeling to create the first integrated, modular computational modeling framework for cancer biology. This will mean that it will no longer be necessary to "re-invent the wheel" as the next generation of researchers rewrite code that was written by previous researchers. It will provide the flexibility that is absolutely

Manuscript received May 14, 2010.

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

Jonathan Cooper is with the Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford, OX1 3QD, United Kingdom. (e-mail: jonathan.cooper@comlab.ox.ac.uk).

Steve McKeever is with the Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford, OX1 3QD, United Kingdom. (e-mail: steve.mckeever@comlab.ox.ac.uk).

essential in allowing the modeller to focus on key issues for a particular application. Moreover, tool support in the form of type and unit checking will enable structural integrity of all models [10].

A. Code Construction

A construction phase will automatically translate the models into optimized code for simulation on computational frameworks that support the required numerical solvers. Tailoring the compilation will be achieved through appropriate construction metadata. This is data that is distinct from the models but describes how we would like them to be executed (SED-ML [11] could be incorporated here, although we will need more features than it currently provides). The construction need not be one to one. In fact a single description could be used to create multiple simulation codes, where each is capable of addressing a specific scientific issue based not only on a set of initial parameters but also on the selection of appropriate components. We also aim to provide a run-time framework similar to what Chaste (see below) provides for the heart modeling community but developed in such a way as to ensure low coupling. Other platforms will rapidly be supported to ensure portability of models.

A second family of metadata will target compile-time analysis techniques by describing the parameters of the overall system and those of each individual component. A generic platform based on SBML/CellML/FieldML allows for generic optimizations. Both continuous and discrete analysis methods will be supported. Various sensitivity analysis techniques will be studied to check for robustness or for allowing more coarse-grained components to be switched in and out during run-time calculations. This would occur based on the accuracy required to ensure computational tractability. A separate class of analysis techniques would rely on traditional compilation methods and examine the types of parameters to ensure more efficient evaluation using provably correct optimizations [12], [13]. Developing such generic optimization techniques means that they could be applied to all models in our repository. It is also important to ensure that these optimizations are correct, in other words that they do not introduce errors into our calculations.

B. The Chaste Framework

We plan to apply the proven functionality of the Chaste [14] computational framework, developed for cardiac electrophysiology, to the area of multi-scale cancer modeling. Both application areas make use of a common core for describing meshes, linear algebra (based on PETSc [15]) and solving ordinary differential equations and partial differential equations (using the finite element method). On the cardiac side, CellML is used to describe the single cell models, which are systems of ordinary differential equations. The PyCml tool [16] translates these CellML descriptions into Chaste-compatible C++ code. Currently, Chaste can load (tetrahedral) meshes in several formats, including VTK and the 'triangle' data format. It is envisaged that once FieldML and a suitable API exist, support for this standard will be added to Chaste in a similar fashion, albeit that FieldML provides more information than just a mesh. Support for SBML is being considered, and also for CellML 1.2 in the future. As a starting point, we will begin with the encoding of a continuous model in a joint CellML and FieldML specification. This encoding will be used to drive our construction phase development that will create a prototype compiler for such applications. A second phase would focus on discrete model formulations (cellular automata and agent-based modeling) that will be described using the state features of the emerging CellML 1.2 format. Once this second compilation phase is complete we would aim to integrate all three cell based models, each addressing a different problem, on to one platform.

For coupling multi-scale models a third family of metadata will be developed. An ontology is needed to describe the environment for each component and not just the parameters. Thus we plan to include contextual interactions, such as the adhesive properties of the membrane or the extra-cellular environment in which the tumors reside, into our models. The long-term goal is to allow modellers to retrieve models from different repositories and couple them together in order to run simulations that address a specific scientific query in a seamless fashion.

Our approach will be incremental and allow the bio-chemistry to be developed alongside the coupling of other phenomena. We aim to provide robust tool support in order to facilitate mathematical modeling and ensure a greater degree of modularity and integrity of resultant simulation codes.

IV. Conclusions

Portability of computational cancer models will be essential for facilitating future research in an international research environment. In this paper we described our aim to extend on emerging model markup languages that have developed out of cardiac modeling, to apply them to models used in cancer modeling. Enabling portability of cancer models, along with the development of authoring and code construction tools, will serve to facilitate cooperation, sharing, and advancement in *in silico* oncology.

REFERENCES

[1] M. Tsiknakis et al, "Semantic Grid Infrastructure Enabling Integrated Access and Analysis of Multilevel Biomedical Data in Support of Postgenomic Clinical Trials on Cancer," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 2, pp. 205-217, March 2008.

[2] K. Marias et al, "Contra Cancrum Website." Internet: www.contracancrum.eu [May 14, 2010].

[3] T.S. Deisboeck, L. Zhang, and S. Martin, "Advancing Cancer Systems Biology: Introducing the Center for the Development of a Virtual Tumor, CViT," *Cancer Informatics*, vol. 5, pp. 1-8, March 2007.

[4] C.M.Lloyd, M.D.B.Halstead, and P.F.Nielsen, "CellML: its future, present and past," *Progress in Biophysics and Molecular Biology*, vol. 85, pp. 433-450, 2004.

[5] 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, June 2004.

[6] G. Richard Christie and P.M.F. Nielsen, S.A. Blackett, C.P. Bradley and P.J. Hunter, "FieldML: Concepts and Implementation," *Philosophical Transactions of the Royal Society A.*, vol. 367, no. 1895, pp. 1869-1884, May 2009.

[7] D. Gianni, S. McKeever, and N. Smith, "euHeartDB: A Web-enabled Database for Geometrical Models of the Heart," in *Proc. 5th Int. Conf. Functional Imaging and Modeling of the Heart*, Nice, France, LNCS 5528, pp. 407-416, 2009.

[8] A. Garny et al, "CellML and Associated Tools and Techniques," *Philosophical Transactions of the Royal Society A.*, vol. 366, no. 1878, pp. 3017-3043, September 2008.

[9] Y. Sun, S. McKeever, K. Balali-Mood, and M. Sansom, "A Multiscale Model for Efficient Simulation of a Membrane Bound Viral Fusion Peptide," in *Proc. IEEE Int. Conf. Bioinformatics and Biomedicine*, Fremont, CA, 2007, pp. 294-301.

[10] J. Cooper, and S. McKeever, "A model-driven approach to automatic conversion of physical units," *Software - Practice and Experience*, vol. 38, no. 4, pp. 337-359, April 2008.

[11] D. Köhn and N. Le Novère, "SED-ML An XML Format for the Implementation of the MIASE Guidelines," in *Proc. 6th Int. Conf. Computational Methods in Systems Biology*, Rostock, Germany, LNBI 5307, pp. 176-190, 2008.

[12] J. Cooper, and S. McKeever, "Experience report: a Haskell interpreter for CellML," in *Proc. 12th ACM SIGPLAN Int. Conf. Functional Programming*, Freiburg, Germany, 2007, pp. 247-250.

[13] J. Cooper, S. McKeever, and A. Garny, "On the Application of Partial Evaluation to the Optimization of Cardiac Electrophysiological Simulations," in *Proc. ACM SIGPLAN Symposium on Partial Evaluation and Semantics-based Program Manipulation*, Charleston, SC, 2006, pp. 12-20.

[14] J. Pitt-Francis et al, "Chaste: a Test-driven Approach to Software Development for Biological Modeling," *Computer Physics Communications*, vol. 180, no. 12, pp. 2452-2471, December 2009.

[15] S. Balay et al, "PETSc Web page." Internet: www.mcs.anl.gov/petsc [May 14, 2010].

[16] J. Cooper, "Automatic Validation and Optimisation of Biological Models," D.Phil. thesis, University of Oxford, Oxford, UK, 2009.

A collaborative system for the *in silico* oncology domain: Requirements, solutions and guidelines

Ioanna Lykourentzou, Dimitra Dionysiou and George Stamatakos

***Abstract*— *In silico* oncology is a demanding research field that entails the development of complex simulation models for the prediction of malignant tumor growth and the response of normal tissue to therapeutic modalities. A system that would enable researchers exchange, share and validate their models could therefore aid the *in silico* domain to the ultimate benefit of cancer patients. This paper presents the requirements of such a system, the solutions that can be implemented to address them, as well as a number of guidelines for its successful implementation.**

I. Introduction

In Silico Oncology is a research field dedicated to the simulation of malignant tumor growth and to the response of normal tissue to therapeutic modalities at all levels of biocomplexity. To perform the above, researchers use clinical data combined with multiscale, computer simulation tumor models. The significant amount of data used, the complexity of the developed models, the effort required to develop them, as well as the need for qualitative and accurate model building, indicate the need for a system that would enable researchers to exchange and share their models, to build on each other's findings, progressing, in this way, the domain to the greater benefit of cancer patients.

Web 2.0 technologies, with the collaborative, knowledge sharing and participative environments that they offer, can be used in this context to facilitate the aforementioned exchange, collaborative editing and dissemination of *in silico* oncology models and data.

This paper presents the requirements that a collaborative system for the *in silico* field poses, followed by the respective solutions that can be used to address them. Additionally, a number of guidelines for the successful implementation of such a system are proposed, drawn out of the collaborative systems' literature with respect to the specific problem.

The rest of this paper is organized as follows: section 2 provides a brief overview of the general philosophy and the main technologies of web 2.0 collaborative systems. Section 3 analyses the requirements of the system and the potential solutions to address them, while it also presents a list of guidelines for its successful implementation. Finally section 4 concludes the paper.

Manuscript received May 5, 2010.

I. Lykourentzou is with the School of Electrical and Computer Engineering, National Technical University of Athens, Greece (phone: +30-210-772-2538; fax: +30-210-772-2538; e-mail: ioanna@medialab.ntua.gr)

D. Dionysiou is with the School of Electrical and Computer Engineering, National Technical University of Athens, Greece (e-mail: dimdio@esd.ece.ntua.gr)

G. Stamatakos is with the School of Electrical and Computer Engineering, National Technical University of Athens, Greece (e-mail: gestam@central.ntua.gr)

II. Web 2.0 collaborative technologies

Web 2.0 technologies [3] have been attracting a growing interest in recent years, due to their ability to facilitate web user interaction, collaboration and innovation creation. They include various types of technologies, such as wikis, blogs, RSS feeds, folksonomies and mashups. The openness, simplicity and support offered by these systems at a low cost, has led a number of user communities, from various different contexts, including educational, health, corporate and governmental ones, to increasingly invest in their adoption. The value that Web 2.0 technologies have gained these diverse fields indicates that they can also be particularly helpful in facilitating collaboration and innovation creation inside specific research fields, such as the *in silico* oncology domain.

III. A collaborative system for the *In Silico* Oncology research domain

A. Requirements and solutions

Overall, the system should facilitate the collection, exchange and validation of cancer models, data and *in silico* modeling scenarios. To do so, a number of issues need to be addressed.

1) *Collaboration enabling:* To facilitate collaboration among researchers, a number of the aforementioned Web 2.0 technologies can be examined and customized to meet the specific needs of the in silico domain. For instance, the system could include co-authoring features, to improve scientific documentation of the stored models and methodologies, folksonomies to enable collaborative content categorization, syndication technologies – such as RSS – to notify users regarding content changes, visualization techniques, to provide users with workspace

awareness, as well as rating and commenting functionalities over the stored models and methodologies to enable their assessment and optimization by the user community.

2) *Information organization:* Although the open editing nature of a collaborative system is likely to facilitate knowledge sharing and innovation creation, yet it is also likely to produce unstructured knowledge content. It is therefore necessary to provide structure support, in order to facilitate navigation and querying, as well as inserting and retrieving useful information. Full compatibility between the different modeling approaches (top-down, bottom-up) used in the *in silico* domain should also be supported, to enable a unified model storage, search and exchange. Taking into account the above, possible organization approaches could include simple taxonomies, standard fixed ontologies or time-varying ontologies which change as they are being used by the repository users to reflect their changing organization needs.

3) *Quality assurance:* Quality assurance is important in the *in silico* research setting in a two-fold manner: it can help cultivate a notion of trust among researchers - that their content will not be misused- and it can also ensure that one will find qualitative information that can indeed add value to their work. To this end, advanced methods of automatic quality assurance found in the literature can be extended and used. As an example in [1], quality tags are assigned by system users to each uploaded model through a peer review process and the aggregated quality result is compared to the threshold set by the community. Then, if a specific model is found to be of inadequate quality, an automatic mechanism is triggered to identify the individuals that can enhance it.

4) Personalization and usability: Since a large number of users with different needs, backgrounds and interests are expected to use the system, personalization and usability are two more important aspects to consider. To enable personalization, different views based on different types of users (basic science researchers, clinicians, biomedical engineering industrial developers) can be implemented. User-friendliness can be achieved through graphical interfaces that facilitate the use of the system or provide statistical analyses over various aspects of the models.

5) Security and privacy: Ensuring main security and privacy requirements (such as confidentiality, authenticity, integrity and non-repudiation) is crucial to cultivate a sense of trust among users and help maintain the level of exchanged information high and on-topic. To enable secure and reliable system operation, efficient cryptographic mechanisms and schemas (membership criteria, scope of valid IP addresses, versioning capabilities) can be adopted.

B. Guidelines towards a successful collaborative in silico system launch

The following guidelines are drawn out of the collaborative system literature with respect to the above identified specific system needs.

1) *Cultivate knowledge sharing spirit:* A system such as the above, no matter how well designed, will not earn the value it was intends to, if it is not used by researchers and stakeholders. Thus, to ensure user participation it is critical to firstly establish a knowledge sharing culture among the community. This can be performed by providing users with specific incentives for participation [2], such as authorship and reviewer quality recognition, as well as by enhancing team spirit through allowing users to formulate their own communities of practice inside the general system.

2) Pre-populate the system: Another means of supporting user participation is to pre-populate the system, prior to its launch, with existing content, since users are less likely to see future value and contribute to an entirely blank platform. The pre-population content can be retrieved through the solutions previously shared by the community, including databases, e-mail communications, forum discussions and other documentation repositories.

3) Provide effective training: User training is a highly significant factor to the success of collaborative systems. Training should be provided regarding both the cultural and the technological aspects of the *in silico* collaborative platform. As far as the cultural aspect is concerned, training should focus on introducing users to the necessary knowledge sharing culture, so that they will be willing to share their models and ideas with colleagues and build the open environment necessary for the system to create value for the community. Technical support is also important and it can be provided through the use of guidelines, tutorials, help pages and training areas inside the system where users can better get accustomed to the platform.

4) *Observe, maintain and evaluate:* Finally, it is also important to constantly monitor, assess and maintain the results achieved, to ensure that the system will be a viable, effective solution, which will produce long-term value for the research community. To this end, dedicated users can be assigned with the task of maintaining the inserted information quality, while visualization techniques combined with techniques such as social network analysis can be implemented to analyze the activity levels inside the system and provide an insight to the collaboration aspects that need to be supported.

IV. CONCLUSION

This paper presents the requirements that a collaborative system, targeted at the specific needs of the *in silico* research domain, has. For each requirement a number of potential solutions, drawn out of the web 2.0 literature, are suggested. Finally, a list of guidelines, for the successful implementation of such a system, is also proposed.

REFERENCES

[1] I. Lykourentzou, K. Papadaki, D. J. Vergados, D. Polemi, V. Loumos, "CorpWiki: A self-regulating wiki to promote corporate collective intelligence through expert peer matching", *Information Sciences*, vol. 180(1), Special Issue on Collective Intelligence, pp. 18-38, 2010.

[2] T. W. Malone, R. Laubacher and C. Dellarocas, "Harnessing Crowds: Mapping the Genome of Collective Intelligence", *CCI Working Paper, 2009-001,* MIT Center for Collective Intelligence, 2009.

[3] T. O'Reilly, "What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software", *Communications & Strategies, No. 1,* p. 17, 2007.