

Sinergy: how semantic can improve early prevention of skin cancers

Diletta Romana Cacciagrano¹✉, Flavio Corradini¹, Leonardo Vito², Laura Cavalieri³

¹Science and Technology, Computer Science Division, University of Camerino, Italy

²Researcher funded by L.I.L.T., Gagliole, Italy

³University of Camerino - ADiTech s.r.l., Ancona, Italy

Received 17 September 2013; Accepted 19 September 2013; Published 14 October 2013

Competing interests: the authors have declared that no competing interests exist.

Abstract

Melanoma is the major source of deaths related to skin cancer. The project "Sinergy", in cooperation with Lega Italiana Lotta Tumori (L.I.L.T.) and ADiTech s.r.l., aims at realising a Cloud-based and semantic-driven framework for screening and analysis of skin lesions. The framework aims at being an ubiquitous laboratory in Cloud, equipped with semantic-driven tools for making analysis, defining/executing/sharing in-silico experiments, mining laboratory data and activities. The engine kernel, a multi-scale and multi-physics skin lesion ontology, not only makes intuitive any operation for non-expert users, but also improves diagnostic processes thanks to the possibility of globally reasoning on skin lesion variables relative to different biological scales (e.g., genetic, molecular, cellular, tissutal ones) and to non-biological ones (e.g., age, gender, location, geography, race).

Motivation and Objectives

Cancer is a class of diseases characterised by out-of-control cell growth and tissue invasion.

Melanoma is the major source of deaths (75%, with lifetime risk for Caucasians of one in 39 for men and one in 58 for women) related to skin cancer.

When melanoma progresses to metastatic stage, powerful mechanisms to resist chemotherapy, radiation and biological intervention are established in the neoplastic lesions, thus hampering the efficacy of current medical therapies and drugs (Helmbach *et al.*, 2001).

Because advanced skin cancers remain incurable, early prevention is mandatory. Research focused on genetic level, diagnosis based on conventional screening tests and treatment by surgical excision are currently the only approaches to reduce mortality.

Research

Since cancer initiation seems to depend on a series of genetic mutations affecting intrinsic cellular programs, the vast majority of cancer research to date has focused on the identification of these genetic and molecular properties of cancer cells. Two highly related genes, KN2A and CDK4, were discovered to harbour germline mutations in roughly 50% of melanoma pedigrees.

However, genetic predisposition can only be found in 3% of all cases. It means that the inheritance of CDKN2A and CDK4 mutations looks insufficient to lead to melanoma in all carriers.

Empirical studies confirm such a conjecture, suggesting that the focus of the research has to necessarily include also other parameters, related to (1) other biological scales (not only the genetic one) and to (2) other physical system (not only the biological one). A model is considered to be multi-scale if it spans two or more different spatial scales and/or includes processes that occur at two or more temporal scales. It is considered to be multi-physics if it involves multiple physical models or multiple simultaneous physical phenomena

For what concerns (1), consider the angiogenesis (Ziemys *et al.*, 2011), a significant transforming phase in tumour growth. Drugs delivered to tissues will not only change the behaviour of melanoma cells (secretion of cytokines, proliferation, differentiation, apoptosis, or migration) in the intracellular drug-triggered cell division process, but also inhibit the development of new capillary sprouts by preventing sprouts from receiving vascular endothelial growth factors. In turn, inadequate glucose and oxygen transported from the blood vessel will drive even more melanoma cells towards apoptosis.

The crossed synergy among organ, tissue and cell scales in angiogenesis hardly interferes on drug distribution and therapeutic effects (i.e., on molecular scale). For instance, several molecular drugs developed to treat melanoma cancer did not work as well *in vivo* as *in vitro* because of absorption, distribution, metabolism, or toxicity problems (Soengas and Lowe, 2003). Many synergistic drug delivery methods have been devel-

oped to increase the drug effect *in vivo*, but it is difficult to quantitatively evaluate their performance. All that highlights the need of defining and exploiting models and methodologies for studying skin melanoma cancer at different biological scales.

Many mathematical models have been proposed to address the challenge mentioned above. These models studied one or more phases of cancer progression, including tumour growth, angiogenesis, and drug treatment, with the purpose of better understanding the pathophysiology of cancer, mechanisms of drug resistance and the optimisation of treatment strategies. Although biologists have already obtained many experimental data sets at the molecular, cellular, micro-environmental and tissue levels, only a few scientists have integrated these data into a n-scale tumor model, where n is greater than two (e.g., atomic-molecular (Liu *et al.*, 2007), molecular-microscopic (Athale and Deisboeck, 2006), microscopic-macroscopic (Zheng *et al.*, 2005)).

For what concerns (2), several heterogeneous factors (e.g., age, gender, location, geography, race) seem to play an important role in melanoma predisposition, incidence and distribution. In some way, such factors should be also related to biological variables (and to genetic ones, in particular). However, it cannot be formally proved yet. In detail:

- **Age:** the relationship between the incidence of melanoma and age is unusual in comparison to other common cancers. There is not an exponential increase in risk with age but rather a more even distribution across age groups.
- **Gender:** men are more likely than women to develop melanoma (67% higher incidence) and their prognosis is worse (136% higher risk of death from melanoma).
- **Location:** basal cell cancers arise exclusively from cutaneous sites and are closely related to sites of skin that receive the most sun exposure, such as the scalp, face, neck, and arms. A small percentage of melanomas arise on acral surfaces of the hands and feet, which tend to be diagnosed at a later stage.
- **Geography:** the rates of melanoma and other skin cancers are highest where fair-skinned Caucasians migrated to lower latitudes.
- **Race:** Caucasians are by far the most susceptible race for melanomas. Hispanics

have a lower incidence but represent the group at next highest risk. Asians and African-Americans have the lowest rates of skin cancer. All racial groups are equally likely to develop melanoma on the acral surfaces of the hands and feet or mucosal surfaces.

Diagnosis and treatment

Conventional screening tests are based on a naked-eye examination by an experienced clinician. ABCD (Soyer *et al.*, 2004) - looking at asymmetry of the skin lesion, irregular edges (borders), color variegation, and diameter - is one of the easiest and most widely used algorithm for evaluating suspicious pigmented skin lesions (PSLs). Since a naked-eye ABCD may fail to detect many borderline PSLs that are small or/and regular in shape or colour, the algorithm is often combined to dermoscopy - a diagnostic technique for an *in vivo* better visualisation of PSLs - and/or implemented in computer-aided image analysis systems (CalASs).

Dermoscopy improves diagnostic sensitivity by 20–30% compared with a naked-eye diagnosis (Soyer *et al.*, 2004). However the results of dermoscopic examination have limitations, especially for the inexperienced, and they are effective only if the user is formally trained.

CalASs allow PSLs to be classified by automatically quantifying ABCD attributes and other texture-based characteristics. However, CalASs effectiveness depends largely on dataset, feature selection and classification methods used (Boldrick *et al.*, 2007). Consequently, conventional CalASs, although pivoting on objective parameters, cannot allow more reliable decision processes than the ones of an experienced clinician (pivoting on experience, complex inferences and extensive knowledge).

“Putting semantic” into CalASs could enable more reliable decision processes, extraction of implicit information from data, integration of quantitative and functional information to infer relevant new knowledge, as well as standardisation of terminology, verification of data consistency and integration of heterogeneous biomedical databases.

Methods

The project “Sinergy”, in cooperation with Lega Italiana Lotta Tumori (L.I.L.T.) and ADiTech s.r.l., started in March 2013, aims at realising a Cloud-based and semantic-driven framework for

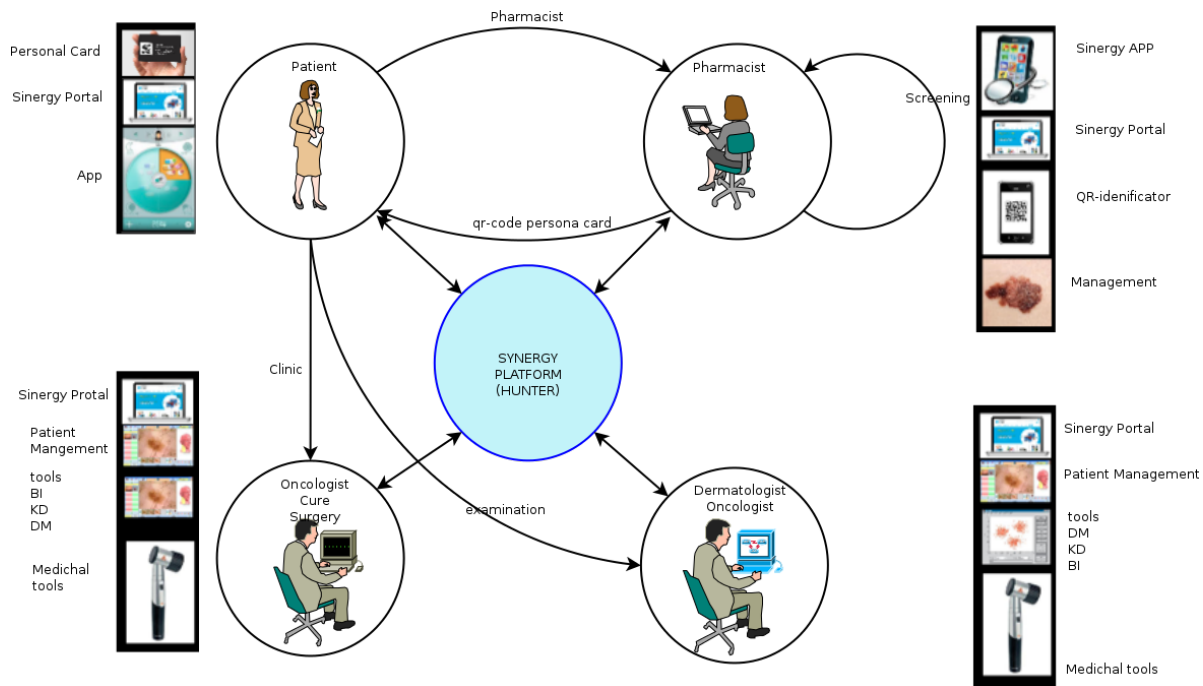


Figure 1. A Sinergy Platform scenario.

screening and analysis of skin lesions. The framework is conceived as the integration of two existing prototypes, UBiolab (Bartocci *et al.*, 2012) and OWL-meaning (Cacciagrano *et al.*, 2012), following the architectural schema described in (Bartocci *et al.*, 2007). In detail, the framework:

- is conceived as a semantic Enterprise Resource Planning (ERP) platform, equipped with (i) a semantic knowledge base storing semantically-annotated information, and (ii) semantic-driven services for managing, analysing, meaning, querying and clustering stored information;
- provides a semantic-driven interface for managing and semantically-annotate the knowledge base, as well as for programming, executing and storing (in the knowledge base as a declarative knowledge) workflow-based procedural knowledge (e.g., business processes, in-silico experiments, new services). This feature, in particular, makes possible mining operations on heterogeneous types of knowledge (e.g., it is possible to reason on in-silico experiments altogether with the involved resources);
- relies on a Cloud-based middleware providing the needed integration mechanism between services and knowledge base and

among services themselves. This allows also further services (developed by the framework programming interface) to be plugged in the framework without changing its current architecture.

The core of the framework is a smart knowledge model (Cannata *et al.*, 2005) that allows at contextualising resources w.r.t. a given domain, and activities w.r.t. given resources, making the framework domain-independent. Selecting a specific domain ontology from a repository suffices to customise the framework for a specific domain (that one conceptualised by the loaded ontology).

Moving information from people to doctor (and not the vice versa) is the first step to spread screening facilities and to considerably improve the early detection of skin cancers.

This is the reason why a smartphone-based screening service has been already developed to be plugged into the framework. It enables pharmacists to acquire skin images and to store them in the (remote) knowledge base using conventional smartphones equipped with cheap dermoscopic cameras.

As shown in Figure 1, different actors can participate at the process of data entry and analysis. Pharmacists are driven in the rule of data

entry by the screening service. Data are annotated (on the basis of a suitable multi-scale and multi-physics domain ontology loaded from the framework repository) by the semantic annotation service.

Once information of patients and acquired images are stored on the knowledge base, dermatologists and oncologists can (1) visualise them by the knowledge management service; (2) use image filter services; (3) classify skin lesions by the clustering service (that, differently from a conventional ABCD, can mix images features with different type of information, like age, social aspect, job, location, family etc.); (4) plan altogether a response and, if necessary, call a specific patient for the frontal examination.

Results and Discussion

Currently, we are testing the screening service over one hundred patients, with the collaboration of L.I.L.T. voluntaries and ASP - Azienda Pubblica di Servizi alla Persona "A. Chierichetti" of Gagliole (Macerata, Italy).

For what concerns the Dermatology domain (with focus on melanoma cancer and skin lesions), defining and engineering a semantic multi-scale and multi-physics skin lesion ontology is crucial.

At the moment, the domain ontology takes into account only demographic and social concepts linked to a FMA Ontology subgraph.

To improve the domain conceptualisation, we plan to extend the basis integrating other well-known standard biomedical vocabulary and ontologies. In detail, XML-based medical standards like HL7/CDA, standards for medical image acquisition like DICOM, nomenclature for describing dermatological disorders as PROVOKE, ONTODerm and DermLex, vocabulary for describing diseases as ICD (version 9 and 10) and SNOMED CT, ontologies like UMLS and Gene Ontology.

Acknowledgements

This work has been supported by Lega Italiana Lotta Tumori (L.I.L.T.) grant - Project Sinergy (2013).

References

- Aithale C, Deisboeck TS (2006) The effects of EGF-receptor density on multiscale tumor growth patterns. *J Theor Biol* **238**(4), 771–779. doi:10.1016/j.jtbi.2005.06.029.
- Bartocci E, Cacciagrano D, *et al.* (2007) An Agent-based Multilayer Architecture for Bioinformatics Grids. *IEEE Trans Nanobioscience* **6**(2), 142–148. doi:10.1371/journal.pcbi.0010076.
- Bartocci E, Cacciagrano D, *et al.* (2012) UBioLab: a web-LABoratory for Ubiquitous in-silico experiments. *J Integr Bioinform* **9**(1), 192. doi:10.2390/biecoll-jib-2012-192.
- Boldrick JC, Layton CJ, *et al.* (2007) Evaluation of digital dermoscopy in a pigmented lesion clinic: clinician versus computer assessment of malignancy risk. *J Am Acad Dermatol* **56**(3), 417–421. doi:10.1016/j.jaad.2006.08.033.
- Cacciagrano D, Merelli E, *et al.* (2012) Semantics on the Cloud: toward an Ubiquitous Business Intelligence 2.0 ERP Desktop. *Proceedings of the Sixth International Conference on Advances in Semantic Processing 2012 (SEMAYPRO 2012)* 23–28 September 2012, Barcelona, Spain. Cacciagrano D, Dini P (eds). Curran Associates Inc., Red Hook, NY, USA. Pp. 42–47.
- Cannata N, Merelli E, Altman RB (2005) Time to organize the bioinformatics resourceome. *PLoS Comput Biol* **1**(7), e76. doi:10.1371/journal.pcbi.0010076.
- Helmbach H, Rossmann E, *et al.* (2001) Drug-resistance in human melanoma. *Int J Cancer* **93**(5), 617–622. doi:10.1002/ijc.1378.
- Liu Y, Purvis J, *et al.* (2007) A multiscale computational approach to dissect early events in the Erb family receptor mediated activation, differential signaling, and relevance to oncogenic transformations. *Ann Biom Eng* **35**(6), 1012–1025. doi:10.1007/s10439-006-9251-0.
- Soengas MS, Lowe SW (2003) Apoptosis and melanoma chemoresistance. *Oncogene* **22**(20), 3138–3151. doi:10.1038/sj.onc.1206454.
- Soyer HP, Argenziano G, *et al.* (2004) Three-point checklist of dermoscopy. A new screening method for early detection of melanoma. *Dermatology* **208**(1), 27–31. doi:10.1159/000075042.
- Ziems A, Kojic M, *et al.* (2011) Hierarchical modeling of diffusive transport through nanochannels by coupling molecular dynamics with finite element method. *J Comput Phys* **230**(14), 5722–5731. doi:10.1016/j.jcp.2011.03.054.
- Zheng X, Wise SM, Cristini V (2005) Nonlinear simulation of tumor necrosis, neo-vascularization and tissue invasion via an adaptive finite-element/level-set method. *Bull Math Biol* **67**(2), 211–259. doi:10.1016/j.bulm.2004.08.001.