Perspective of the Large Databases and Ontologic Models of Creation of Preclinical and Clinical Results

18

Elisa Meldolesi, Mario Balducci, Silvia Chiesa, Andrea Damiani, Nicola Dinapoli, Roberto Gatta, and Vincenzo Valentini

General Introduction

Over the past decade, remarkable advances in the medical field, and in particular, in cancer care have occurred, leading to a tremendous transformation in the internal medical concept [1]. Starting from an inflexible "one size fits all similar groups" approach, where the same treatment is used for the same kind of tumor, clinical practice is moving towards a personalized medicine concept with an essential role of decision support systems (DSS).

Glioblastoma multiforme (GM) is the most common primary brain tumor with only few available therapies providing significant improvement in survival. Therefore, the development of new diagnostic and treatment technologies beside the concomitant research progress in pathology, biologic biomarkers (e.g., MGMT promoter, DNA metilation, IDH, EGFR, etc. [2]), genomics, and proteomics justifies the growing trend towards "individualized medicine".

The use and role of medical imaging technologies in clinical oncology has also greatly expanded during the last decade from a primarily diagnostic and qualitative tool to award a central role in the context of individualized medicine with a quantitative value. Several studies have been developed to analyze and quantify different imaging features (e.g., descriptors of intensity distribution, spatial relationships between the various intensity levels, texture heterogeneity patterns, descriptors of shape, etc.) and the relations of the tumor with the surrounding tissues to identify their possible relationship with treatment outcomes or gene expressions [3, 4].

Furthermore, multidisciplinary management of cancer patients has been proven essential to reach a highly individualized treatment. The integration between different specialists leads to a mortality reduction not only cancer-related, but also related to concomitant diseases [5, 6].

In this context of progressive technologies and treatment innovation, the development of predictive models can answer to the increasing necessity of individualized medicine. Based on individual patient features, in fact, predictive models, complementing existing consensus or guidelines, allow physicians deliver tailored treatment. Patient care is transforming from an evidence-based treatment into a personalized medicine concept (build on an evidence base) going from prescription by consensus to prescription by numbers.

E. Meldolesi (✉) • M. Balducci • S. Chiesa
• N. Dinapoli • R. Gatta • V. Valentini
Radiation Oncology Department, Gemelli-ART,
Università Cattolica S. Cuore, Largo A. Gemelli 8,
Rome 00168, Italy
e-mail: elimlds@gmail.com;
mario.balducci.1955@gmail.com;
silvia.chiesa.md@gmail.com;
nicola.dinapoli.74@gmail.com;
roberto.gatta.bs@alice.it; vvalentini@rm.unicatt.it

Personalized Medicine

Personalized medicine is defined by the National Cancer Institute as "a form of medicine that uses information about a person's genes, proteins, and environment to prevent, diagnose, and treat disease. In cancer, personalized medicine uses specific information about a person's tumor to help diagnose, plan treatment, find out how well treatment is working, or make a prognosis" [7].

To date, in the medical field and inherently also in oncology, clinical practice is based on evidence-based guidelines and protocols as results of the outcome of randomized clinical trials (RCTs). Although in the past decades they have had a key role in the definition of the treatment strategies in cancer care, RTCs's population is often constituted by a selective group of patients, very different from the population seen in routine clinical practice. Some patients groups are under-represented, including elderly, those with comorbidities [8, 9], and patients from under-represented ethnic and socioeconomic backgrounds [10–12]. Furthermore, the long time that it is usually requested to reach the pre-established outcome is an intrinsic limitation of this kind of research. As a result, the presented evidence is often valid for only a subgroup of patients and trial results are quickly outdated.

Beside RCTs, a complementary form of research is progressively emerging that has, in the population-based observational studies, its major expression. The role of this new research is mostly to ensure that the result of clinical trials translates into tangible benefits in the general population [13]. Given the differences between patients recruited to trials and those seen in routine practice, in fact, small benefit observed in highly selected trial patients is likely to disappear when the same treatments are applied in routine practice. Observational studies are essential to identify whether practice has changed appropriately, to document harms of therapy in a wider population, in patients of different age and with different comorbidities, and to determine whether patients in routine practice are reaching the expected outcomes with the expected toxicity [14, 15].

In this new era of individualized medicine, it is more and more important to develop supporting decision tool, based on models able to predict different outcomes starting from large heterogeneous datasets. Essential, for the development of this kind of DSS, is the creation of large databases, archives of heterogeneous data coming from multiple sources. Numerous information that are routinely collected in clinical practice as diagnostic and clinical imaging, laboratory data, treatment outcome data, biologic environment, genomics, and proteomics are included into large databases. Using innovative "rapid-learning" research techniques, these data are simultaneously analyzed in order to obtain, from the extraction of knowledge of the masses, a benefit of the individual [16]. From a technical point of view, this large amount of data required to create a predictive model is necessary not only to provide sufficient statistical power to act as an efficient and reliable predictive tool, but also to validate the obtained model. Therefore, a secondary dataset is needed for validation of the model, preferably by external (from a different institution) datasets [17]. Only after external validation, a prediction model can be implemented as an acceptable decision support tool.

In this context, the idea of research is totally changed. Heterogeneity of data is now assuming a key role against the ab initio definition of the collecting variables (as in the RCTs). Large databases approach requires gathering data without knowing beforehand what would be the outcomes of the research, which is quite different from the fixed design of a prospective randomized controlled trial. Therefore, a flexible strategy for data collection, data mining, and outcome reporting is needed with the possibility to add new variables to the large databases in an ad-hoc manner.

Considering that large database can be created combining data coming from various departments of a single hospital or from multiple institutes different on a regional, national, and international level, integration of information is a big challenge for data-sharing initiatives.

Ontology and Data Standardization

The standardization process, essential to universally define data and procedures that will constitute a large database, is obtained through the creation of an ontology.

"Ontology" is a compound word, composed of onto-, from the Greek ὄντος (òntos) which is the present participle of the verb εἰμί (eimi), i.e., "to be, I am", and λόγία (lògia), i.e., "science, study, theory". Ontology formally represents knowledge as a set of concepts within a domain and the relationships between those concepts. In practice, an ontology is a terminological system where all the information, related in this case to medical disciplines and treatment, are specified and organized in a well-defined data collection model. An ontology collects uniform and unambiguous definition for each variable and the relationship between different variables into the space and the time concept. Eventually, better and unambiguous understanding leads to an approach where the research data could be made available without differences in interpretation; for now and the future. From the perspective of computer science, different kind of data can be represented in any ontology starting from a generic "registry" layer with purely epidemiologic information, to a "procedural" level, where treatment information and related toxicities are reported, up to a higher "research" level where dimensional data, such as images, genomics, proteomics, etc., are collected [18]. Therefore, in the development of an ontology, the information can grow both in terms of variety and granularity, until the idea of clinical large database [18].

Furthermore, the formalization of any ontology can grow from a simple dictionary, where the meaning of the terms is described in natural language, toward a more and more formal expression resulting also from the sharing of the definitions between different institutions on a local, national, or international level. At the cost of increasing complexity and formalism that enriches the language with more and more complex constructs representing relationships between variables, different techniques can be used for representing richer knowledge contents. In this context, the most frequently used model to represent data distribution is the Semantic Web, developed by Tim Berners-Lee [19]. For the Semantic Web technology, data is represented by triplets (subject, predicate, object) using the Resource Description Framework (RDF) language [20].

The interaction between elements of multiple triplets is defined inside an ontology through a different language (RDFS or OWL) allowing informatics system to automatically generate inference from any exploitable data source. Software agents can easily parse and make inference on big data repositories applying formal-ontologies on explicitly declared facts to infer the entire set of facts logically inferable.

The power of the semantic web is the extremely simple, however flexible RDF representation (one table with three columns) (Table 18.1), as well as the federated nature of the web where both data and knowledge can reside at multiple locations on the internet and can be queried using SPARQL, the query language of the Semantic Web [21].

Furthermore, a distributed learning approach is able to learn from the collected data creating a

Table 18.1 Examples of "semantic" triple representation [18]

Subject	Predicate	Object example (URL)	Reference
Patient	hasBeenDiagnosedWith	Malignant neoplasm of rectum, http://purl.bioontology.org/ontology/ICD10/C19	ICD-10
Patient	hasBiologicalSex	Male, http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C20197	NCI Thesaurus
		Female, http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C16576	
Disease	hasStageFinding	T1 Stage finding, http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C48720	NCI Thesaurus
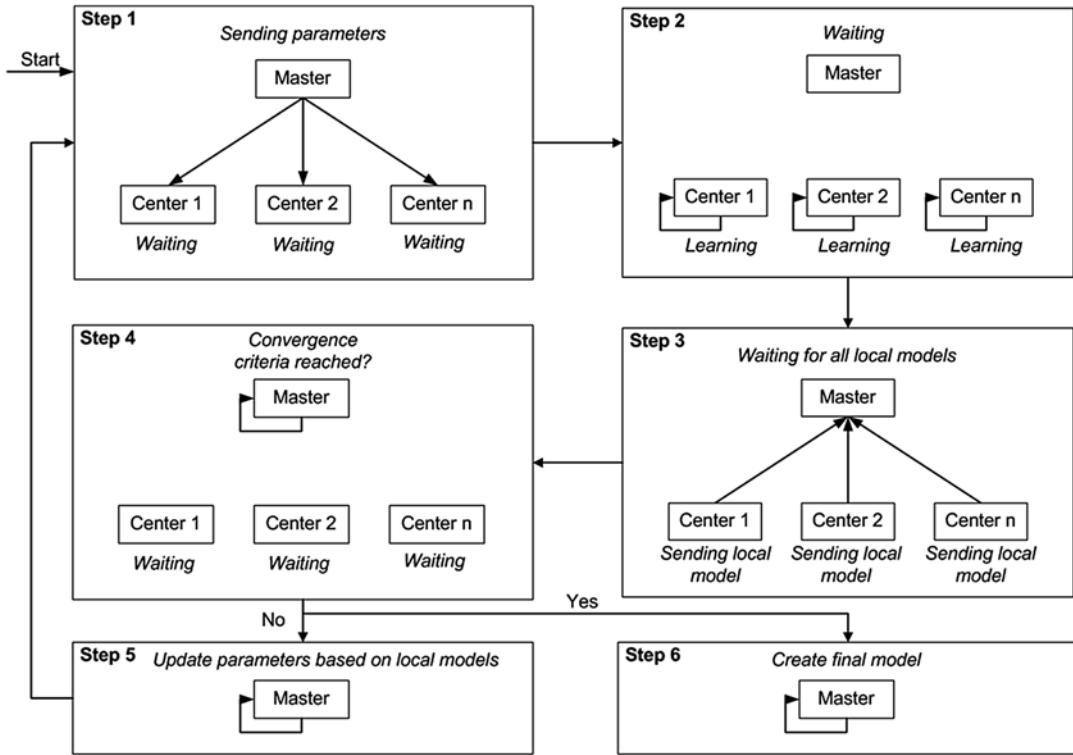
Fig 18.1 Distributed machine learning flow [23]

model without the need for data to leave the individual hospital. A distributed machine learning algorithm is able, through a local learning application which is installed at each hospital, to create a local model that is sent to the central server. Starting from the integration of all the single models, a consensus model is generated and sent back to each hospital for refinement. After pre-established convergence criteria are met, it is possible to create a final consensus model (Fig. 18.1). This method works for a variety of models as described in literature [22].

Radiomics and Imaging Analysis

In medical field and inherently also in oncology, the imaging technologies have always had a key role in the identification and staging of a cancer disease being fundamental for the definition of the treatment procedure. During the last decade, we have witnessed an important change of the medical imaging concept coming from a diagnostic, qualitative position to award a central role in the context of individualized medicine with the identification of numerous measurable features.

The term "Radiomics" is a relatively new term that was used in several studies to indicate the extraction of large amounts of features from radiographic images with the intent of creating mineable databases [3]. The goal of Radiomics is to convert images into mineable data, with high fidelity and high throughput [4].

Until last decade, texture heterogeneity, characteristics of shape, volume, and intensity distribution of the tumor, were only analyzable, on the acquired images, in a qualitative way. In this new Radiomics era, images are fractionated in order to identify specific patterns and/or descriptor that could be quantified and easily reproduced in a consistent manner in different institutions.

Considering the different gray scales inside the tumor image, it was possible to identify and quantify not only some descriptors (e.g., descriptors

of shape, texture, and optical porosity), but also the relationship between the tumor and the surrounding tissues in a bidimensional and tridimensional way [3, 4].

Despite all this technological progress, it is still a long way to identify the numerous heterogeneity's patterns characteristics of different tumors. However, it is clear how these patterns could highly contribute to choose the better treatment strategy for each single patient.

Prediction Models

Over the past decade, medical doctor had to face numerous and remarkable challenges in oncology that have progressively moved toward a personalization of the treatments. In this context of growing technologies and treatment's innovation, predictive models achieve a relevant role, beside the existing consensus and/or guidelines, in helping clinicians in daily clinical practice.

The methodological process to develop a DSS is depicted in Fig. 18.2 [1].

A large heterogeneous database is required to store all the information without knowing beforehand what would be the research's topic. From the hypothesis, it is determined which features should be included in the learning effort. Bayesian network is usually considered the best approach [24] to impute for the missing data and to detect and correct bias into the initial dataset, to improve data quality. After this pre-processing step, it is possible, through a machine learning procedure, to analyze the different features listed in the large database and obtain a model representing the distribution of the same features and their relationship inside the dataset.

Beside common medical statistics approaches (Cox proportional hazard model [25], logistic regression [26] etc), the usage of different machine learning algorithms (Bayesian network [27, 28], decision trees [29], support vector machines [30], neural networks [31], genetic algorithm [32], etc.) leads to the possibility of creating predictors characterized by different performance and usage related to the final outcome. To obtain a reliable and consistent DSS and able to work properly also in a different environment from where it was created, it is necessary to validate the new model (training set) preferably by external dataset (validation set) [1, 17].

Considering the performance, the Receiving Operating Characteristic (ROC) and its equivalent Area Under the Curve (AUC) are the most used measurement units (Fig. 18.3). However, it is important to know that the ROC is not always applicable to all the predictor: in such cases different indicators could be used (accuracy, sensitivity, specificity, F-score, etc.).

To date, European Organization for Research and Treatment Cancer (EORTC) has developed several interactive DSS related to either primary or recurrent glioblastoma (Table 18.2). These survival's prediction models are currently used in clinical practice beside the existing consensus and/or guidelines, helping clinicians in choosing the better treatment strategy for each single patient.

Medical doctors and/or patients can use predictive models in a variety of ways. Graphical calculating devices as nomograms [25, 33] are one of the most common forms of predictive device, beside the even more appealing interactive website (Table 18.2). Furthermore, in this era of technological progress, the possibility to create specific applications for devices of new generation is also very interesting (e.g., cell-phones, tablet, etc.).

Perspectives in Glioblastoma

GM is the most common primary brain tumor, but, even now, only few available therapies providing significant improvement in survival are known. In the past decade, the possibility to use more and more sophisticated technologies allowed to deal with numerous challenges obtaining a tremendous influx of data describing molecular and genomic alterations in the pathogenesis of GM [34]. Notwithstanding this explosion of knowledge, the early clinical data from the usage of selective therapies developed on these identified aberrations are largely disappointing. The wide heterogeneous nature of this disease and the possibility for the tumor to change mutations during its progression, beside the well-known difficulty of neuro-oncology
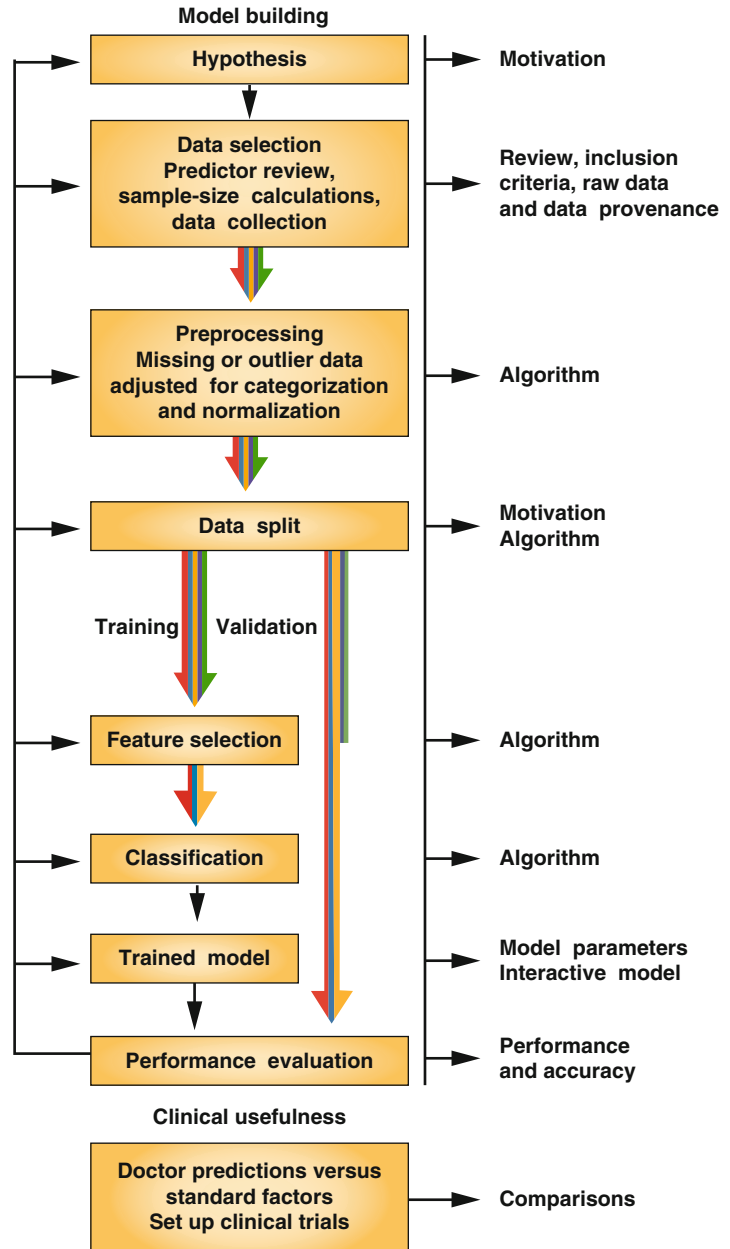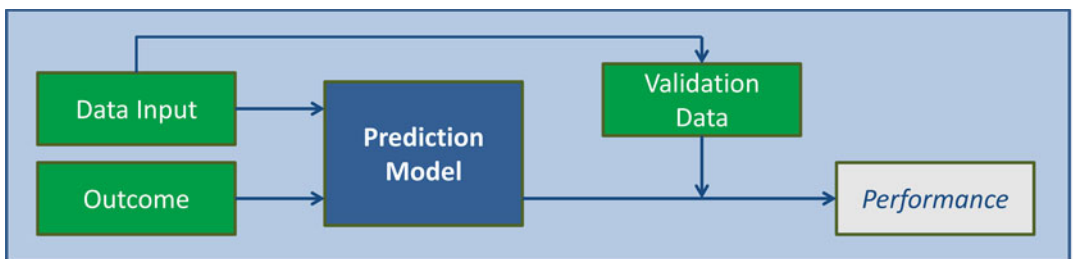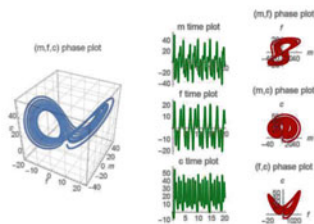
Fig 18.2

Figure 1| Schematic overview of methodological processes in clinical decision-support system devolopment, describing model development, assessment of clinical usefulness and what ideally to publish. The coloured, parallel lines represent heterogeneous. The which have been split early for independent validation (but without internal cross-validation).

Table 18.2 Examples of interactive decisional support systems (DSS) related to glioblastoma, currently used in clinical practice

| Institution	Prediction model	Web link (URL)
European Organisation for Research and Treatment of Cancer (EORTC)	Prediction of survival in general GMB population	https://www.eortc.be/tools/gbmcalculator/model1.aspx
European Organisation for Research and Treatment of Cancer (EORTC)	Prediction of survival in patients treated by RT/TMZ (MGMT methylation status unavailable)	https://www.eortc.be/tools/gbmcalculator/model2.aspx
European Organisation for Research and Treatment of Cancer (EORTC)	Prediction of survival in patients treated by RT/TMZ (MGMT methylation status available)	https://www.eortc.be/tools/gbmcalculator/model3.aspx
European Organisation for Research and Treatment of Cancer (EORTC)	Prediction of survival in patients with recurrent glioblastoma	http://www.eortc.be/tools/recgbmcalculator/calculator.aspx
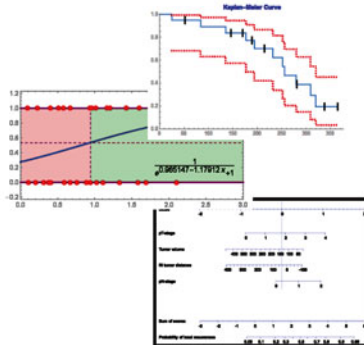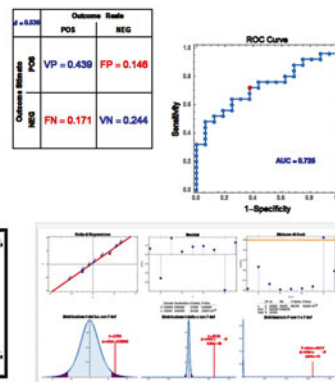
Fig 18.3 Predictive models

drugs to penetrate the blood–brain barrier, can partially justify the large ineffectiveness of the most current molecular-targeted therapies. Despite these discouraging initial results, it is still very reasonable to believe that in the era of "individualized medicine" genomically and molecularly driven research in combination with multiple patients-specific data (clinical, pathological, biological, proteomics, imaging, etc.) will ultimately be successful.

Recent studies have demonstrated how the interaction between an imaging's quantitative analysis and specific gene and microRNA tumor

expression can be useful as a robust initial prognostic tool in order to personalize therapy for GBM patients [35, 36]. Therefore, only through the understanding of the gene regulatory network and the study of the interaction between molecular alteration and different GM's characteristic features, it will be possible to develop better preclinical models that will help physicians to choose the best drug or the best combination of drugs for each patient in the most efficient possible way.

Conclusions

The interaction between the implementation of new technologies and the usage of automated computer bots has allowed, in the last decade, a broad range of researches to be expanded, due to the very generalizable and flexible technology utilized. In oncology, the availability of reliable and consistent prediction tools makes possible to stratify population in specific risk groups for different selected outcomes, identifying patients who better than other can benefit from a specific treatment procedure. Furthermore, it will also stimulate research focused on specific risk groups, trying to find new treatment options or other combinations of treatment options for these subgroups. Therefore, personalized medicine can be expected not only to save patients from unnecessary toxicity and inconvenience, but also to facilitate the choice of the most appropriate treatment.

Clinicians are now facing two new challenges. The first one is represented by the trend towards "individualized medicine" trying to consider several potential options for each patient in place of inflexible "one size fits all similar groups" approach. Secondly, the new concept of "prescription by numbers" support the moving towards a "shared decision making" approach, where doctors and patients, evaluating pros and cons of different treatment strategy, can actively discuss and decide on therapeutic interventions.

The development and validation of predictive models is a fundamental step to create new software able to give the knowledge a different dimension. Guidelines and protocols currently used in a daily clinical practice will be optimized by the usage of predictive models, considering that medical doctor will have a more accurate idea of the treatment's possibilities for each patient in terms of both survival and side effects.

The behavior of specific tumor is very difficult to predict due to their huge intrinsic heterogeneity. However, treatment can only become more personalized if accurate, science-based decision aids are developed, which can offer assistance in clinical decision-making in daily practice.

Therefore, the poor human cognitive capacity, able to discriminate and use not more than 5 features in a daily clinical practice [37], can find in DSS a valuable help able to compensate for this human intrinsic limitation.

Finally, considering the important role that predictive models could play in the clinical practice, clinicians must be aware that although they can be very useful with great performances and sometimes with a great p-value, they remain only DSS, not decision-makers.

References

1. Lambin P, van Stiphout RGPM, Starmans MHW, Rios-Velazquez E, Nalbantov G, Aerts HJWL, et al. Predicting outcomes in radiation oncology—multifactorial decision support systems. Nat Rev Clin Oncol. 2013;10(1):27–40. http://www.ncbi.nlm.nih.gov/pubmed/23165123.
2. McNamara MG, Sahebjam S, Mason WP. Emerging biomarkers in glioblastoma. Cancers (Basel). 2013;5(3):1103–19. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3795381&tool=pmcentrez&rendertype=abstract.
3. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48(4):441–6. http://www.ncbi.nlm.nih.gov/pubmed/22257792.
4. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. Magn Reson Imaging. 2012;30(9):1234–48.
5. Mason WP, Del Maestro R, Eisenstat D, Forsyth P, Fulton D, Laperrière N, et al. Canadian recommendations for the treatment of glioblastoma multiforme. Curr Oncol. 2007;14(3):110–7. http://www.ncbi.nlm.nih.gov/pubmed/17593983.
6. Friedland PL, Bozic B, Dewar J, Kuan R, Meyer C, Phillips M. Impact of multidisciplinary team management in head and neck cancer patients. Br J Cancer.

2011;104(8):1246–8. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3078600&tool=pmcentrez&rendertype=abstract.

7. NCI. NCI Dictionary of terms: personalized medicine. http://www.cancer.gov/dictionary?cdrid=561717. http://www.cancer.gov/dictionary?cdrid=561717

8. Tyldesley S, Zhang-Salomons J, Groome PA, Zhou S, Schulze K, Paszat LF, et al. Association between age and the utilization of radiotherapy in Ontario. Int J Radiat Oncol Biol Phys. 2000;47(2):469–80. http://www.ncbi.nlm.nih.gov/pubmed/10802375.

9. Faivre J, Lemmens VEPP, Quipourt V. Bouvier a M. Management and survival of colorectal cancer in the elderly in population-based studies. Eur J Cancer. 2007;43(15):2279–84. http://www.ncbi.nlm.nih.gov/pubmed/17904353.

10. Bach PB, Cramer LD, Warren JLBC. Racial differences in the treatment of early-stage lung cancer. N Engl J Med. 1999;341(16):1198. http://www.nejm.org/doi/full/10.1056/NEJM199910143411606.

11. Boyd C, Zhang-Salomons JY, Groome PA, Mackillop WJ. Associations between community income and cancer survival in Ontario, Canada, and the United States. J Clin Oncol. 1999;17(7):2244–55. http://www.ncbi.nlm.nih.gov/pubmed/10561282.

12. Hershman D, McBride R, Jacobson JS, Lamerato L, Roberts K, Grann VR, et al. Racial disparities in treatment and survival among women with early-stage breast cancer. J Clin Oncol. 2005;23(27):6639–46. http://www.ncbi.nlm.nih.gov/pubmed/16170171.

13. Booth CM, Tannock IF. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. Br J Cancer Nature Publishing Group. 2014;110(3):551–5. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3915111&tool=pmcentrez&rendertype=abstract.

14. Booth CM. Evaluating patient-centered outcomes in the randomized controlled trial and beyond: informing the future with lessons from the past. Clin Cancer Res. 2010;16(24):5963–71. http://www.ncbi.nlm.nih.gov/pubmed/21169249.

15. Fuentes-Raspall R, Puig-Vives M, Guerra-Prio S, Perez-Bueno F, Marcos-Gragera R. Population-based survival analyses of central nervous system tumors from 1994 to 2008. An up-dated study in the temozolomide-era. Cancer Epidemiol. 2014;38(3):244–7. http://www.ncbi.nlm.nih.gov/pubmed/24794586.

16. Lambin P, Roelofs E, Reymen B, Velazquez ER, Buijsen J, Zegers CML, et al. 'Rapid Learning health care in oncology'—an approach towards decision support systems enabling customised radiotherapy. Radiother Oncol. 2013;109(1):159–64. http://www.ncbi.nlm.nih.gov/pubmed/23993399.

17. Vickers AJ. Prediction models: revolutionary in principle, but do they do more good than harm? J Clin Oncol. 2011;29(22):2951–2. from: http://www.ncbi.nlm.nih.gov/pubmed/21690474.

18. Meldolesi E, van Soest J, Dinapoli N, Dekker A, Damiani A, Gambacorta MA, et al. An umbrella protocol for standardized data collection (SDC) in rectal cancer: a prospective uniform naming and procedure convention to support personalized medicine. Radiother Oncol. 2014;112(1):59–62. http://www.ncbi.nlm.nih.gov/pubmed/24853366.

19. Tim B-L, Hendler James LO. "The Semantic Web" A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. Sci Am Mag. 2001;284:35–43. http://www.cs.umd.edu/~golbeck/LBSC690/SemanticWeb.html.

20. Graham Klyne JJC. Resource Description Framework (RDF): concepts and abstract syntax. 2004. http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/.

21. Assélé Kama A, Choquet R, Mels G, Daniel C, Charlet J, Jaulent M-C. An ontological approach for the exploitation of clinical data. Stud Health Technol Inform. 2013;192:142–6. http://www.ncbi.nlm.nih.gov/pubmed/23920532.

22. Boyd S. Distributed Optimization and Statistical Learning via the alternating direction method of multipliers. Found Trends® Mach Learn. 2010;3(1):1–122. http://www.nowpublishers.com/product.aspx?product=MAL&doi=2200000016

23. Meldolesi E, van Soest J, Alitto AR, Autorino R, Dinapoli N, Dekker A, et al. VATE: VAlidation of high TEchnology based on large database analysis by learning machine. Color Cancer. 2014;3(5):435–50. http://www.futuremedicine.com/doi/abs/10.2217/crc.14.34.

24. Luta G, Ford MB, Bondy M, Shields PG, Stamey JD. Bayesian sensitivity analysis methods to evaluate bias due to misclassification and missing data using informative priors and external validation data. Cancer Epidemiol. 2013;37(2):121–6. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3834354&tool=pmcentrez&rendertype=abstract.

25. Valentini V, van Stiphout RGPM, Lammering G, Gambacorta MA, Barba MC, Bebenek M, et al. Nomograms for predicting local recurrence, distant metastases, and overall survival for patients with locally advanced rectal cancer on the basis of European randomized clinical trials. J Clin Oncol. 2011;29(23):3163–72. http://www.ncbi.nlm.nih.gov/pubmed/21747092.

26. Bagley SC, White H, Golomb BA. Logistic regression in the medical literature. J Clin Epidemiol. 2001;54(10):979–85. http://linkinghub.elsevier.com/retrieve/pii/S0895435601003729.

27. Jayasurya K, Fung G, Yu S, Dehing-Oberije C, De Ruysscher D, Hope A, et al. Comparison of Bayesian network and support vector machine models for two-year survival prediction in lung cancer patients treated with radiotherapy. Med Phys. 2010;37(4):1401–7. http://www.ncbi.nlm.nih.gov/pubmed/20443461.

28. Oh JH, Craft J, Al Lozi R, Vaidya M, Meng Y, Deasy JO, et al. A Bayesian network approach for modeling local failure in lung cancer. Phys Med Biol. 2011;56(6):1635–51. http://www.ncbi.nlm.nih.gov/pubmed/21335651.

29. Hu R. Medical Data Mining Based on Decision Tree Algorithm. Comput Inf Sci. 2011;4(5):14–9. http://www.ccsenet.org/journal/index.php/cis/article/view/11950.

30. Van Stiphout RGPM, Lammering G, Buijsen J, Janssen MHM, Gambacorta MA, Slagmolen P, et al. Development and external validation of a predictive model for pathological complete response of rectal cancer patients including sequential PET-CT imaging. Radiother Oncol. 2011;98(1):126–33. http://www.ncbi.nlm.nih.gov/pubmed/21176986.

31. Patel JL, Goyal RK. Applications of artificial neural networks in medical science. Curr Clin Pharmacol. 2007;2(3):217–26. http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1574-8847&volume=2&issue=3&spage=217.

32. Wiggins M, Saad A, Litt B, Vachtsevanos G. Genetic Algorithm-Evolved Bayesian Network Classifier for Medical Applications.

33. Gorlia T, van den Bent MJ, Hegi ME, Mirimanoff RO, Weller M, Cairncross JG, et al. Nomograms for predicting survival of patients with newly diagnosed glioblastoma: prognostic factor analysis of EORTC and NCIC trial 26981-22981/CE.3. Lancet Oncol. 2008;9(1):29–38. http://www.ncbi.nlm.nih.gov/pubmed/18082451.

34. Bastien JIL, McNeill KA, Fine HA. Molecular characterizations of glioblastoma, targeted therapy, and clinical results to date. Cancer. 2015;121(4):502–16.

35. Zinn PO, Mahajan B, Majadan B, Sathyan P, Singh SK, Majumder S, et al. Radiogenomic mapping of edema/cellular invasion MRI-phenotypes in glioblastoma multiforme. PLoS One. 2011;6(10), e25451. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3187774&tool=pmcentrez&rendertype=abstract.

36. Zinn PO, Sathyan P, Mahajan B, Bruyere J, Hegi M, Majumder S, et al. A novel volume-age-KPS (VAK) glioblastoma classification identifies a prognostic cognate microRNA-gene signature. PLoS One. 2012;7(8), e41522. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3411674&tool=pmcentrez&rendertype=abstract.

37. Abernethy AP, Etheredge LM, Ganz PA, Wallace P, German RR, Neti C, et al. Rapid-learning system for cancer care. J Clin Oncol. 2010;28(27):4268–74. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2953977&tool=pmcentrez&rendertype=abstract.