Current Pharmacology Reports (2018) 4:145–156
https://doi.org/10.1007/s40495-018-0127-4

PRECISION MEDICINE AND PHARMACOGENOMICS (S NAIR, SECTION EDITOR)

# Data-Driven Methods for Advancing Precision Oncology

Prema Nedungadi[1,2] · Akshay Iyer[1] · Georg Gutjahr[1] · Jasmine Bhaskar[1,2] · Asha B. Pillai[3]

Published online: 6 March 2018


## Abstract

**Purpose of Review** This article discusses the advances, methods, challenges, and future directions of data-driven methods in advancing precision oncology for biomedical research, drug discovery, clinical research, and practice.

**Recent Findings** Precision oncology provides individually tailored cancer treatment by considering an individual's genetic makeup, clinical, environmental, social, and lifestyle information. Challenges include voluminous, heterogeneous, and disparate data generated by different technologies with multiple modalities such as Omics, electronic health records, clinical registries and repositories, medical imaging, demographics, wearables, and sensors. Statistical and machine learning methods have been continuously adapting to the ever-increasing size and complexity of data. Precision Oncology supportive analytics have improved turnaround time in biomarker discovery and time-to-application of new and repurposed drugs. Precision oncology additionally seeks to identify target patient populations based on genomic alterations that are sensitive or resistant to conventional or experimental treatments. Predictive models have been developed for cancer progression and survivorship, drug sensitivity and resistance, and identification of the most suitable combination treatments for individual patient scenarios. In the future, clinical decision support systems need to be revamped to better incorporate knowledge from precision oncology, thus enabling clinical practitioners to provide precision cancer care.

**Summary** Open Omics datasets, machine learning algorithms, and predictive models have enabled the advancement of precision oncology. Clinical decision support systems with integrated electronic health record and Omics data are needed to provide data-driven recommendations to assist clinicians in disease prevention, early identification, and individualized treatment. Additionally, as cancer is a constantly evolving disorder, clinical decision systems will need to be continually updated based on more recent knowledge and datasets.

**Keywords** Precision oncology · Precision medicine · Health analytics · Predictive analytics · Artificial intelligence · Big data in health · Personalized medicine · Omics · Clinical decision support

## Introduction

Cancer is a multifaceted disease, driven by selected modifications of genes and proteins at both the genetic and epigenetic levels. Nearly one in six deaths was due to cancer in 2015, making it one of the leading causes of mortality worldwide. According to the WHO, the incidence of cancer is expected to increase by 70% over the next two decades. Common subtypes of cancer are lung cancer, liver cancer, colorectal cancer, stomach cancer, and breast cancer [1].

Traditional methods of treating cancer include chemotherapy, hormone therapy, and biological modifiers such as cancer growth factor inhibitors, radiation therapy, and surgery. Traditional treatments for cancer, which are based on type,

This article is part of the Topical Collection on *Precision Medicine and Pharmacogenomics*

✉ Prema Nedungadi
prema@amrita.edu

[1] Center for Research in Analytics & Technology in Education, Amrita Vishwa Vidyapeetham, Amritapuri, Kollam, India

[2] Department of Computer Science, School of Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, Kollam, India

[3] Division of Pediatric Hematology/Oncology, Departments of Pediatrics and Microbiology and Immunology, University of Miami Miller School of Medicine, Miami, FL, USA

stage, and histological grade, are suboptimal due to large interperson variability in response and toxicity profiles.

Precision oncology (PO) is a paradigm shift from treatment plans based on expected outcome of the average patient to treatment plans tailored for disease prevention and maximum therapeutic benefit for each individual patient.

PO identifies mechanisms of cancer evolution, develops therapies for subgroups of patients, and supports physicians to provide targeted patient care in a timely manner. For example, molecular testing has been used to identify genetic markers, such as EGFR for lung cancer, which helps identify the most effective medication to be administered [2, 3]. It integrates omics-based data with clinical, environmental, social, and behavioral information to identify subpopulations of patients based on genomic alterations and thus greatly improves the probability of successful prevention or treatment.

**Cancer Prevention and Early Detection** PO aims to enhance cancer prevention through targeted risk reduction models and also to improve survivorship through early detection. Risk prediction models may use genetic status along with environmental and behavioral information to gauge the risk of developing disease. In addition, identifying cancer at an early stage significantly improves survivorship. In the case of breast cancer [4, 5], for example, liquid biopsies detect tumor DNA in the blood stream and offer the possibility of detecting cancer with a blood test much before cancer is visible. [6, 7]

**Drug Development and Therapy** In traditional medicine, development of a cancer drug requires such information as oncogene addiction, cancer cell specificity, tumor response, and therapeutic impact [8]. Even in the case of same type of cancer, treatment responses of drugs vary. Modeling is performed with the assistance of software such as BIO-CAD or BODIL [9, 10]. Precision medicine follows the concept that the effectiveness of the drug depends on the patient's individual genetic profile, and therefore, a particular drug can have varying levels of effectiveness over different patients. Thus, PO adds another level of complexity by studying the patient's specific genetic profile as well as the nature of the cancer to determine the most effective treatment for an individual [2].

**Cancer Treatment** Patients with a similar cancer subtype often respond differently when treated with the same chemotherapeutics. Recent research has focused on exposing the complex interplay of genomics and chemotherapeutic sensitivity, resistance, and toxicity [11]. Therapies based on precision medicine depend on the patient's genotype, phenotype, and biomarkers. Identifying biomarkers to help determine the most appropriate treatment for cancer offers many benefits including higher probability of selecting an effective initial treatment and avoidance of treatments likely to be ineffective [12].

Targeted therapies act on mutations that cause cancer progression by turning certain genes on or off. Compared to chemotherapy and radiation, targeted therapies often have fewer and less severe side effects as they reduce the "off-target" toxicity to non-tumor cells [13].

**Immunotherapy** Immunotherapy drugs such as monoclonal antibodies, chimeric antigen receptor (CAR)-T cells, and anti-tumor vaccines stimulate or direct the immune system to better recognize and kill tumor cells [14]. Immune checkpoints are a mechanism used by cancer cells to escape from the immune system. To address this problem, immune checkpoint inhibitors have been effective in treating a variety of tumor types including advanced lung cancer and melanoma [15].

**Combination Therapy** Precision oncology has the potential to expand use of combination therapies and explain causes of drug resistance. For patients with higher and more dangerous levels of malignancies, combination therapy involving multiple immunotherapies as well as traditional chemotherapy [16] may be effective. The blockading of immune checkpoints has been one of the most powerful breakthroughs in cancer immunotherapy, with the intent of eliciting the anti-tumor specific T cell response [17–20].

### Data Analytics

Data analytics has propelled precision oncology in all areas of biomedical research, drug discovery, data-driven clinical trials, predictive analytics, and clinical decision systems. This requires data integration and analysis of heterogeneous and disparate data from multiple modalities (Table 1). Rapid improvements in technologies related to "omics," increasing amounts of data generated, and customized data analytic methods, have exponentially expanded advances in PO. Technologies such as gene therapy and next-generation sequencing are continuing to advance and become more affordable and available to the general public [3].

## Data Sources and Characteristics for Precision Oncology

The simultaneous convergence of several factors has created both challenges and opportunities for precision oncology: the exponential growth of large-scale datasets in omics, clinical trials and cancer imaging, improvements in computational genomics, deep learning and predictive analytics methods,

**Table 1** Diverse data in precision oncology

| Data source | Data types | Characteristics | Examples |
|---|---|---|---|
| Omics | Sequence data | High dimensional, uncommon distributions.<br>Heterogeneous, data noise (technical and biological).<br>Lack of standards for Next-Generation Sequencing (NGS) data | Repositories such as Genbank and Uniprot. HUGO: human genes<br>dbSNP: point mutations |
| HIS | Structured, unstructured, image, text | Longitudinal patient data.<br>Volume—large amounts of records; but individual record sparse.<br>Veracity—noisy, incomplete (more as clinical notes). | EHR |
| Other HR | Structured data | Sparse data | Death registers; cancer register; insurance company records |
| Wearables, sensors and mobile technologies | Streaming data | Continuously monitored data, high throughput<br>Velocity—data accumulated at high speeds. | Blood pressure monitoring, ECC monitoring, blood glucose monitoring, heart rate monitoring, and so on. |
| Clinical trials | Efficacy and toxicity information | Higher standard in data collection, quality and cleaner data. Smaller data. | dose finding, survival trials |
| Other sources | Spatial data, environmental data | Variety | Ecological survey, assessment of water quality, EPA air quality, toxic data, pesticide use. |
| Epidemiological studies | Surveys, aggregated data, lifestyle data | Structured, semi-structured | Dietary studies, studies to estimate prevalence and incidence of disease. |

and the dramatic drop in cost of sequencing a human genome from $100 million in 2001 to $1000 [21].

The large amount of data generated by different technologies with multiple modalities provides continuing challenges for data integration, analysis, and interpretation. The data obtained by the next-generation sequencing is around 3 GB and can vary up to 200 GB. For example, large-scale personal genomics and pharmacogenomics datasets have been generated and are used to uncover unique signaling pattern that occur in specific subgroups of patients. Drugs have been developed that target these unique patterns.

Data analytics for precision oncology broadly consists of (a) biomedical research such as identification of biomarkers and cancer causing gene evolution pathways; (b) drug discovery in repurposing and discovering targeted drugs based on resistance or sensitivity; (c) data-driven clinical trials to help in population stratification, adaptive trials, and identifying patients from electronic health record (EHR) and omics systems for clinical trials; (d) modeling and clinical decision systems that help researchers and clinicians in early detection, diagnosis, and combination therapies (Fig. 1). Primary data sources for precision medicine include the following:

- Patient biological data
- Scientific literature such as Medline and PubMed
- Data obtained directly from clinical trials
- Clinically significant variant database such as ClinVar
- EHR/EMR, death/cancer registries, insurance data
- Wearables/sensors
- Medical Imaging data such as The Cancer Imaging Archive [22]
- Molecular Omics sources such as the Encyclopedia of DNA Elements (ENCODE) [23], Genotype Tissue Expression (GTEx) [24], the NIH Epigenomics Project [25], The Cancer Genome Atlas [26, 27] and Drug Screen [28].

## Types of Data and Data Sources

### Omics and Cancer Databases

Omics databases contain comprehensive catalogs of molecular profiles such as the genomic (DNA), transcriptomics (transcribed RNA from genes), epigenetic (methylation profiles of tissues), proteomic (protein profiles of specific tissues and cells), and metabolic data in biological samples. Together, this information is dubbed "systems biology" and provides a holistic view of the organism [29].

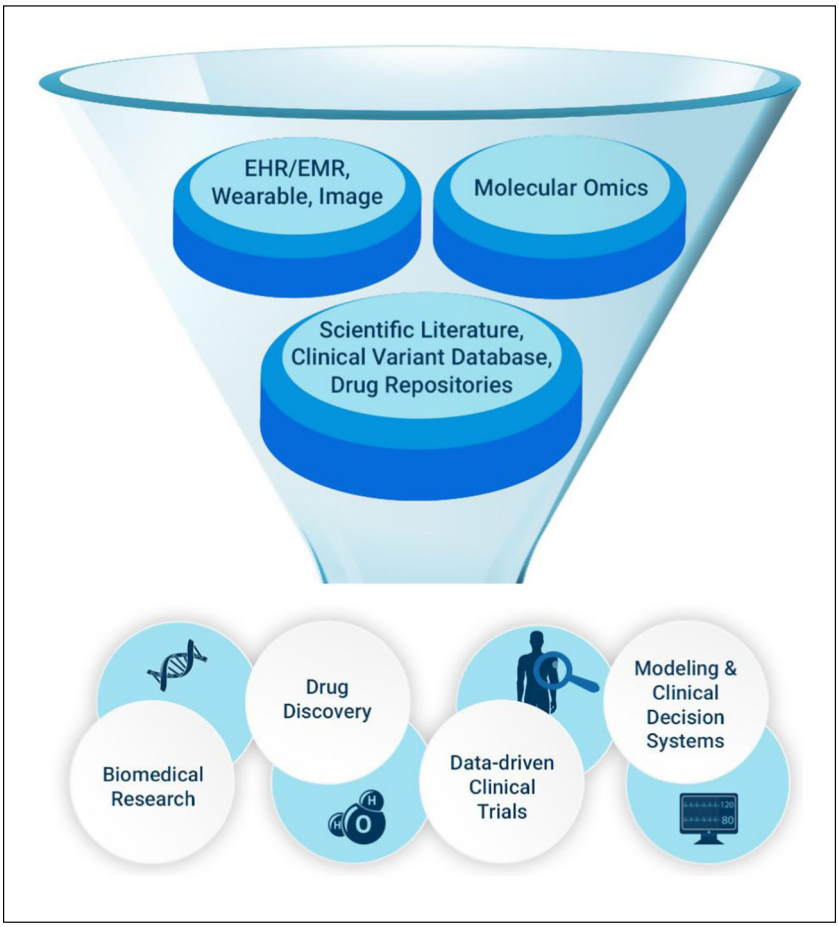


**Fig. 1** Data sources and analytics driving precision oncology: data analytics for precision oncology broadly consists of (a) biomedical research, (b) drug discovery, (c) data-driven clinical trials, and (d) modeling and clinical decision systems. This requires data integration, analysis, modeling, and interpretation of heterogeneous and disparate data from multiple modalities

Different types of Omics data can be found from repositories that include genomic data (such as UCSC Cancer Genomics Browser), DNA methylation (such as MethyCancer), transcriptome data (such as NONCODE), drug sensitivity and response information (such as GDSC), alliteration and mutation based data (such as cBioPortal), and multidisciplinary information (such as canSAR) [30]. Sequence data is generally stored in FASTQ or Bam formats.

The Pan-Cancer project, developed by the Cancer Genome Atlas research network, aims to analyze multiple tumor types and molecular aberrations in cancer types to enable scientists to discover new aberrations [26]. Similarly, several projects such as the Cancer Cell Line Encyclopedia [31] and the Genomics of Drug Sensitivity in Cancer [32] are generating large genomic databases to specifically interrogate links between genomic biomarkers and drug sensitivity in cancer cell lines.

Large pharmacogenetic databases can be used to predict drug sensitivity. Computational algorithms for predicting drugs can be improved based on genomic profiles and drug-response data [11]. Based on these predictions, clinical trials may be designed to test the results of such algorithmic predictions on tumor response and toxicity in patients.

## Medical Imaging Data

Quantitative medical imaging provides tumor phenotype data including tumor shape, size, volume, and texture. These features can be correlated with clinical outcomes data and used for evidence-based clinical decision support in conjunction with the other information provided by clinical reports, omics, and lab tests.

Medical imaging data is commonly generated utilizing techniques such as computed tomography (CT), magnetic resonance imaging (MRI), or positron emission tomography (PET). A hybrid scanner that combines the two modalities into a single scan is called MRI/PET. Incorporating the information from both anatomy and metabolic activity helps to better distinguish benign from malignant nodules or masses found in imaging. This knowledge improves disease characterization, treatment evaluation and restaging, and decreases unnecessary radiation exposure to the patient.

Computer-aided detection (CAD) has also been used in cancer imaging. Computer-aided diagnostics (CADx) can measure the malignancy by using the features of the image. New systems proposing an integrated CAD system that can both detect and diagnose nodules in lung cancer have been proposed [33].

Radiomics is the field that studies the processing and analyses of medical tomographic images. Radiomics provides imaging biomarkers with potential to help in detecting and diagnosing cancer, assessment of prognosis, prediction of response to treatment, and monitoring of disease status. The multimodal imaging features in radiomics are useful for predicting prognosis and therapeutic responses [34]. Radiomics supports clinical decision support systems as it extracts a large number of quantitative features from digital images and mines the data for hypothesis generation [35].

The processing and analysis of images in oncology contain the following parts: first, the region of the tumor in the image is segmented; second, feature extraction is used from the tumor region; and third, feature selection is performed based on the goal of the clinical application [36].

The methods used in radiomics include the methods from classical image engineering for the prepossessing and segmentation [37], from machine learning for the feature extraction [38], and from multivariate statistics for individual feature (covariate) selection and interpretation [39].

## Electronic Health Record

EHR are digital records containing the patient medical information such as medical history, patient diagnosis and treatment, laboratory examination, medication, and ancillary clinical data [40].

These records include unstructured data, such as clinical notes, and structured data including International Classification of Diseases (ICD) codes and administrative data. Data from EHR can be used to generate hypotheses about risk factors for cancer and to improve the care for a patient [41], for example, by identifying adverse drug effects based on certain patient characteristics [42]. Furthermore, EHR-linked DNA biobanks allow research in precision medicine [43].

Analyzing data from EHR often includes a large amount of preprocessing and data cleaning [44]. Missing values are very common in EHR, and imputation methods can be used [45]. Since the important parts of EHR are in free-text form, methods from natural language processing are often required before additional analysis can be done [46]. Finally, various data mining techniques, such as association rule learning and sequential pattern learning, are used to extract information of clinical interest [47].

Difficulties in obtaining EHR data can occur due to the patient's lack of follow-up, unreachability due to recovery, treatment at another medical facility, or death [48].

## Clinical Trials

Clinical studies are investigations of patients in a clinical setting. Some examples of the use of clinical studies in PO include the validation of biomarkers [49], finding the correct dosage of a target or immune-activating drug [50] and showing the effectiveness of such a drug [51]. Clinical trials can be distinguished into interventional studies, where patients are randomized to different treatments and observational studies, where the research and researcher collect data on subjects without involving in treatments.

In interventional clinical trials, the investigator has control over the treatment that each patient receives. Conducting such trials requires the approval by health authorities and an ethics commission. The Declaration of Helsinki, currently in its seventh revision, provides the basis for such national regulatory bodies that are responsible for approval of a clinical trial [52]. Different types of clinical trials can be distinguished based on how the patients are randomized to treatments [53]. For example, in precision medicine, the effectiveness of a target drug may be investigated in a parallel group study, where a predefined number of patients with particular molecular alterations are randomized to either a conventional treatment or to the target drug; alternatively, the effectiveness of the target drug could also be investigated by a cluster design, where patients are separated into clusters and subsequently clusters are randomized to different treatments.

In observational studies, randomization is not under the control of the investigator and therefore such studies are susceptible to various forms of bias and confounding. For this reason, they lack the internal validity of and do not provide the same level of evidence as randomized trials [54]. However, observational studies are useful in understanding public health; furthermore, in some cases, such as when it is not ethically feasible to randomize patients into specific interventional groups (toxicity studies, etc.), observational studies are the only ethically appropriate study design [55]. An example of the latter situation in oncology is in an investigation of the abortion–breast cancer hypothesis, a controversial hypothesis that says that abortion increases the risk for breast cancer [56]. This study could not be performed by a randomized or prospective design given the ethical issues around forced abortions.

The most common outcomes in oncological studies are binary data (for example, presence or absence of disease or death), incidence rates for relapse, and progression-free survival times. Common statistical methods to analyze such data include logistic regression and the Cox proportional hazards model [57, 58].

## Sensor Informatics: Wearable, Implantable, and Ambient Sensors

The collection of health information can be performed through use of various technologies. Wearable body sensors continuously measure factors such as pulse, blood pressure, sleeping patterns, and inflammation. This information can help medical professionals improve the quality of treatment. The field of data stream mining is concerned with extracting information from such sensors [59]. Methods used to analyze sensor data

differ from other data analysis techniques in that they analyze sequentially, instead of by processing the data in batches. Several machine learning algorithms for clustering, regression, and classification have been modified to work with these kinds of data [60].

## Precision Oncology Research Areas

Table 2 shows the research focus areas that form the building blocks of Precision Oncology, along with the applications, methods, and outcome.

### Prognostic and Predictive Biomarkers

There are two types of biomarkers: prognostic and predictive. Prognostic biomarkers guide in the determination of oncology outcome risks, while predictive biomarkers help determine effective therapy decisions. Sometimes, a biomarker may have both good prognostic and strong predictive properties (e.g., circulating tumor cells [CTCs]) [71].

The identification of new biomarkers starts with the process of identifying the distinctive features and measuring the quantity of biological molecules that transform or convert into the composition, function, and dynamics of an organism. Biomarkers are commonly analyzed using techniques such as reverse transcriptase or proteomic analysis such as CyTOF® mass spectrometry.

Data is analyzed using a variety of multivariate statistical techniques, depending on the data type. For example, pattern recognition-based techniques such as principal component analysis (PCA) are utilized to screen LC/MS or NMR-based data.

### Gene Evolution Pathway

This type of data is drawn from collections of regulatory molecules, their interactions, and their effect on gene and gene expression in a pathway. After sequencing, these pathways are generally interrogated through comparative studies such as an enrichment or nodal analysis. Gene functions can be derived from this, and the obtained information can be used to generate new drugs and biomarkers [72]. The Cancer

**Table 2** Research Methods in precision oncology

| PO research areas | How it is used | Method | Outcome |
|---|---|---|---|
| Identify prognostic biomarkers and predictive biomarkers. | Facilitate diagnosis.<br>Determine effective therapy decisions. | Association studies between genes, proteins, and cancer onset. | Has shown good results in several cancer types such as RCC [61] or prostate cancer [62] |
| Cancer causing gene evolution pathways | Determine genes influencing the onset of cancer.<br>Develop specific, gene-targeted therapy. [2, 8] | Nodal analysis, NGS techniques, whole genome sequencing, patient studies, pathway interaction studies.<br>Pathway inhibition studies | Understanding of several pathway mechanics that cause cancer onset.<br>Several corresponding risk genes such as KIT, MET, and PDGFRA have been identified from these pathways that help in early identification of cancer [63]. |
| Classify patient population biomarker subgroup selection. | Used to identify patients, and individuals susceptible to cancer.<br>Used to classify subgroups of patients that can respond to specific treatments | Staining, molecular technologies, sequencing, subgroup analysis and multiple testing methods | Recently, biomarkers such as PD-L1 IHC, which relates to immunoblockading, have been of interest [64], as well as PDGFR alpha [65] |
| Developing gene targeted drugs. | Used in "targeted therapy" (targets and influences specific, pre-identified, genes.)<br>Drug resistance studies and prediction models. | Enrichment studies, clinical trials, sequencing studies.<br>Mechanism-based mathematical modeling for drug resistance [66]. | Targets specified gene mutations and alterations. Some studies show increased effectiveness with the incorporation of nanoparticles [67]. |
| Developing immune-activating drugs | Focuses on activating and directing an anti-tumor immune response, generally mediated by T cells [68]. | Analysis of pathways connected to apoptosis (particularly about ligands such as PD-L1), enrichment studies, clinical trials, sequencing studies. | Advances for selective targeting for cancer in the past few years have produced several effective drugs such as nivolumab [69, 70]. |
| Estimating cancer progression | Making decisions about treatment, continuous monitoring, preventive medicine. | Modeling depending on omics, lifestyle, environment, data mining, Markov models. | Individual risk factors, avoiding invasive procedures for patients, cost-effective for the healthcare provider. |

Genome Anatomy Project (https://cgap.nci.nih.gov/) is one of the most comprehensive databases on cancer genomics, transcriptomics, and proteomics. Predictions are made using genome-wide association studies (GWAS) and gene set enrichment analysis (GSEA).

### Drug Research and Therapy

New research is being conducted to understand drug sensitivity based on omics, with the goal of developing less toxic treatments for cancer. In the past few years, research has been invested in the treatment of cancer at the molecular level, leading to new kinds of cancer treatments such as the targeting intracellular signaling pathways.

The programmed death-ligand 1 (PD-L1) pathway has been found to be of particular interest with regard to cancer. Monoclonal antibodies have been developed to the receptors controlling these pathways (CTLA-4, PD-1/PD-L1) and have been quite successful in the past few years [18].

Identification of genetic mutations that can influence the development of cancer has led to development of several drugs which have shown impressive responses in the majority of patients. The effectiveness of these drugs tends to be short lived, so techniques such as combination therapy may be useful in increasing effectiveness of the treatment [19].

Mina et al. designed an algorithm to analyze molecular data from the Cancer Genome Atlas (TCGA) international consortium (6456 genomes from 23 tumor types). This algorithm identified cancer evolutionary dependencies from genomic alteration occurrences. While these genomic alterations were shown to impart resistance to certain drugs, it also resulted in novel discoveries of tumor subtype sensitivity to other drugs [20].

Companion diagnostic tests are used to stratify patients with the FDA providing guidance on co-developing these along with the drug to help identify patients who will benefit the most or with serious adverse effects.

### Data-Driven Clinical Trials

In observational studies, molecular alterations and histology can be treated as covariables, while in interventional clinical trials, they can also be used to define inclusion-exclusion criteria and to perform stratified randomization [73]. Moreover, when molecular profiling is done while recruiting, it is possible to perform a data-driven trial [74].

Methods for such trials include group sequential designs [75], enrichment designs for subpopulation selection [76], adaptive treatment selection based on genetic information [77], subgroup selection based on predictive biomarkers [78], and endpoint selection based on observed effectiveness [79]. The main advantage of such a trial is flexibility [80]; secondary advantages may be an increase in power if the right subgroups or the right endpoint are selected [79] and a potential reduction in sample size in sequential designs [80].

## Modeling and Clinical Systems for Precision Oncology

The development of technologies that produce high-throughput data (such as RNA-seq or mass spectrometry) has made mathematical modeling techniques an absolute requirement for interpretation. Because of the availability of reference data (found in repositories such as GEO or ENCODE), modeling can produce accurate output that can be used for hypothesis-driven studies. The different kinds of modeling approaches used for systems biology can be broadly classified as ordinary differential equation based modeling (ODE), Petri Net-based modeling, Boolean modeling, modeling based on linear programming, and agent-based modeling (AGM) [81].

Ordinary differential equations relate the change in biochemical quantities over time by modeling them as functions and derivatives of functions. In precision medicine, they are primarily used to understand pathway mechanics. Common applications include modeling dynamic gene regulation [82], and calculation of cell growth/death rates [83]. They are also integrated with larger models such as Bayesian Networks which can help analyze large gene regulatory networks [82].

Petri nets are mathematical models that represent a biochemical system as a bipartite graph where nodes represent events and arrows represent preconditions for these events [84]. In precision medicine, Petri nets are used to model complex interactions between gene expression and molecular interactions [85]. Once the system is represented, many properties of the system can be derived automatically [86].

Petri nets are by no means the only formalism that can be used to model such networks of genes. A simple alternative to Petri nets are Boolean network model that use proposition logic to derive properties of the biochemical system, for example genome-wide molecular interactions [87]. While genes are not simple binary switches, it has been observed that the pattern of expressed and suppressed genes can in many cased still often be approximated well by Boolean networks [88, 89]. Another simple alternative is to connect the states of the biological system by linear functions and to use linear regression, linear classification, and linear programming to derive the properties of the system [90]. More complex approach that does not require linear relationships between the components of the system can be analyzed by agent based simulation [91].

With respect to cancer, modeling has been used to analyze the impact of factors such as interactions between the immune system and the environment, or the reaction to therapy.

Mathematical models also include gene-based interactions, drug-based interactions and patient responses, and scans (such as MRI or Xray). This kind of modeling can help construct and address hypotheses, such as the development of resistance by tumor cells [92].

Models researching drug resistance which use mechanisms such as cellular signaling networks as inputs can be broadly classified into mechanistic modeling, which include experimental techniques such as molecular dynamic simulation (used to investigate the development of drug resistance) and data-driven prediction methods such as omics data-based node biomarker screening [66].

### Modeling Cancer Progression

Biological models of tumor progression have various applications and are recognized as promising research tools for oncology [93]. Such models simulate the behavior of individual elements in the cancer treatment paradigm, such as cancer cell responses, immune cell responses, and energy transfer.

### Predictive Models for Cancer Survivorship

Genetic information may be used in predictive models for cancer survival. In parametric and semi-parametric models, such as the Weibull model or the proportional hazards model [94], genetic subgroups can be incorporated as covariables, used for stratified survival models, or treated as a frailty term [95]. Furthermore, penalized and hierarchical survival models have been proposed to deal with the pathway information [96].

### Predictive Analytics for Data-Driven Clinical Trials

Use of predictive analytics (PA) to generate strata reduces the heterogeneity of patients. This may be based on patient treatment outcome or on time-to-disease progression. Strata can be used either up-front in the study design or for post-hoc stratified analysis of data. In identifying strata, PA has the ability to reduce randomization failure and size of clinical trials. It can also decrease bias, provide feedback, and allow drugs to fail earlier in the clinical trial, thus reducing time and cost.

Data mining may be used to understand gene-phenotype and disease relationships and develop disease progression models. For example, PA can be used in identifying FDA-approved dose finding models for clinical trial along with tools to estimate progression of clinical trial patients in the experimental group.

### Clinical Decision Systems to Support Precision Oncology

The decreasing cost of sequencing and enhanced data analysis methods allow for effective analysis of petabytes of biomedical data [97]. According to the 2016 Precision Medicine Essentials Brief, however, only 29% of hospitals in the USA are utilizing precision medicine.

Given the high value, yet complex nature of PM, doctors need to be educated in how to integrate PM into their practices, how to counsel patients based on genomic information, and how to reliably compare patient profiles with possible drugs and intervention in an acceptable amount of time.

Because the field is evolving so quickly, some means of continuing education are required to help doctors keep track of new biomarkers and therapy options as they are being discovered.

Currently, although there are a number of tools for predicting and annotating genomic changes that support variant identification, variant annotation, and visualization, tools to support clinicians in the interpretation of these data in a clinical setting are very limited.

A clinical decision support (CDS) system could ideally include EHR factors such as age, race, gender, and other health parameters such as emergency or long-term care; however, most clinical support systems only use subsets of this data [98]. In rural areas, CDS systems on mobile devices can capture and support community health workers in differential diagnosis and recommending additional diagnostic tests [99]. Quantitative medical imaging features such as intensity, shape, size, volume, and texture can also be included in a CDS. This offers information on tumor phenotype. These features can be correlated with clinical outcomes data used for evidence-based clinical decision support in conjunction with the other information.

Systems that use multiomics data, with algorithms to understand phenotype-genotype, gene-gene and gene-environment interactions, can suggest therapies to clinicians based on predicted drug-efficacy. These approaches are critically needed to facilitate the adoption of PO in clinical settings. This may require a complete overhaul of the cancer treatment ecosystem, with different clinical workflows and personalized information for each patient. The lack of standard interfaces for such integration, along with varying formats used for genomic test, further complicates the process.

## Discussion and Conclusion

The continued emphasis by governments on creating open datasets for omics, drastic reductions in cost of processing genetic information (with an expected drop to $100 within

the next 10 years) [100], along with increased speed and lower cost of high-throughput systems contributing to omics data, and the continuing refinement of mathematical models and predictive analytics has created a unique potential for precision medicine, in general, and precision oncology, in particular, to optimize cancer care delivery. While these factors enable significant new developments in disease treatment, they also present significant challenges. Data integration and analytics must keep up with the exponential increase. A major challenge involves the need to create some means by which to process and disseminate all of this information in meaningful ways to researchers, clinicians, and patients.

PO can help improve cancer therapy by anticipating drug resistance and proposing alternative strategies such as immunotherapy. Genetic alterations in cells pose challenges for precision medicine in the prediction of drug resistance for a tumor cell. Strategies to combat resistance against these drugs must also be taken into account, such as blocking parallel pathways [101]. The evaluation of these genetic aberrations is both time consuming, expensive, and remains yet another challenge to overcome [102].

As genetic profiling becomes more widespread and cost-effective, data-driven designs for clinical trials with increased predictability in outcomes and timelines will become more common. Data-driven patient selection, based on suitability, can optimize the eligibility criteria at the study level [103]. The eligibility criteria can be modified during the trial based on real-time data. Besides the additional financial costs, the main challenges with such trials is the increased complexity of the logistics in performing such trials.

Combination therapies are also commonly in use, allowing a multifaceted approach to cancer. A greater understanding is needed to combat the unexplained drug resistance seen in cancer. PO can help improve cancer therapy by anticipating drug resistance and proposing alternative strategies, such as a shift from chemo/radiotherapy to immunotherapy for an individual patient.

New models of predictive analytics incorporating new knowledge will need to be designed to better predict cancer progression, cancer survivorship, drug discovery, and drug resistance. Extensions of these models have been proposed that incorporate biomarkers [104] and pathway-specific information [105].

In addition to all of the promising research regarding cancer treatment, the data and analytics revolution also holds great potential for cancer prevention. Several important risk factors, such as oncogenes, have been identified and have put us in a position to take preventive measures before the onset of the disease.

Expert clinical and intelligent systems are required to translate knowledge into clinical practice and propel the large-scale adoption of genomic-based precision oncology into clinical practice.

Logistically, it is still difficult to incorporate precision medicine in clinical practice due to lack of integrated systems. In particular, even where implemented, it may take multiple days from the time a biopsy is analyzed until the recommended treatment is found. As time is a critical factor in late-stage cancer, a faster turnaround time is needed to make precision medicine applicable and pragmatic for hospital application.

Many of the existing clinical systems do not assist clinicians in providing PO-based recommendations; hence, patients are unable to benefit from new knowledge in PO for early diagnosis, prevention, or therapies. This is a challenge as hospitals have already invested in a plethora of EHR systems, primarily with the limited goal of hospital and patient management. The lack of standard models for EHR systems, standard interfaces for such integration, and standard representations for genomic test results will continue to challenge interoperability. A redesign of EHR systems to include or even integrate omics data with patient data, ability to update new omics evidence into EHR, along with new workflows, is required to provide real-time and data-driven precision oncology patient care.

In the future, artificial intelligence-based clinical systems will enable clinicians to provide targeted diagnosis and treatment, benefiting cost-effectiveness, outcomes, and the patient's quality of life, while minimizing treatments that may be ineffective or even unnecessarily toxic for that individual. As cancer is a constantly evolving disorder, predictive models and clinical decision support systems will need to continuously update themselves to keep pace with the constant expansion of datasets and new knowledge.

**Acknowledgements** Our work derives inspiration and guidance from the Chancellor of Amrita Vishwa Vidyapeetham, Sri Mata Amritanandamayi Devi.

## Compliance with Ethical Standards

**Conflict of Interest** On behalf of all authors, the corresponding author states that there is no conflict of interest.

**Human and Animal Rights and Informed Content** This article does not contain any studies with human or animal subjects performed by any of the authors.

## References

1. Plummer M, de Martel C, Vignat J, Ferlay J, Bray F, Franceschi S. Global burden of cancers attributable to infections in 2012: a synthetic analysis. Lancet Glob Health. 2016;4(9):e609–16. https://doi.org/10.1016/S2214-109X(16)30143-7.
2. Shin SH, Bode AM, Dong Z. Precision medicine: the foundation of future cancer therapeutics. npj Precision Oncology. 2017;1(1) https://doi.org/10.1038/s41698-017-0016-z.

3. Jameson JL, Longo DL. Precision medicine—personalized, problematic, and promising. N Engl J Med. 2015;372(23):2229–34. https://doi.org/10.1056/NEJMsb1503104.
4. Easton DF, Pharoah PDP, Antoniou AC, Tischkowitz M, Tavtigian SV, Nathanson KL, et al. Gene-panel sequencing and the prediction of breast-cancer risk. N Engl J Med. 2015;372(23):2243–57. https://doi.org/10.1056/NEJMsr1501341.
5. Maas P, Barrdahl M, Joshi AD, Auer PL, Gaudet MM, Milne RL, et al. Breast cancer risk from modifiable and nonmodifiable risk factors among white women in the United States. JAMA Oncology. 2016;2(10):1295–302. https://doi.org/10.1001/jamaoncol.2016.1025.
6. Cheng F, Su L, Qian C. Circulating tumor DNA: a promising biomarker in the liquid biopsy of cancer. Oncotarget. 2016;7(30):48832–41. https://doi.org/10.18632/oncotarget.9453.
7. Esposito A, Criscitiello C, Locatelli M, Milano M, Curigliano G. Liquid biopsies for solid tumors: understanding tumor heterogeneity and real time monitoring of early resistance to targeted therapies. Pharmacol Ther. 2016;157:120–4. https://doi.org/10.1016/j.pharmthera.2015.11.007.
8. Benson JD, Chen YNP, Cornell-Kennon SA, Dorsch M, Kim S, Leszczyniecka M, et al. Validating cancer drug targets. Nature. 2006;441(7092):451–6. https://doi.org/10.1038/nature04873.
9. Sun W, Starly B, Nam J, Darling A. Bio-CAD modeling and its applications in computer-aided tissue engineering. Comput Aided Des. 2005;37(11):1097–114. https://doi.org/10.1016/j.cad.2005.02.002.
10. Lehtonen JV, Still DJ, Rantanen VV, Ekholm J, Björklund D, Iftikhar Z, et al. BODIL: a molecular modeling environment for structure-function analysis and drug design. J Comput Aided Mol Des. 2004;18(6):401–19. https://doi.org/10.1007/s10822-004-3752-4.
11. Sheng Z, Pettersson ME, Honaker CF, Siegel PB, Carlborg Ö. Standing genetic variation as a major contributor to adaptation in the Virginia chicken lines selection experiment. Genome Biol. 2015;16(1):219. https://doi.org/10.1186/s13059-015-0785-z.
12. Khleif SN, Doroshow JH, Hait WN. AACR-FDA-NCI cancer biomarkers collaborative consensus report: advancing the use of biomarkers in cancer drug development. Clin Cancer Res. 2010;16(13):3299–318. https://doi.org/10.1158/1078-0432.CCR-10-0880.
13. Schwaederle M, Zhao M, Lee JJ, Eggermont AM, Schilsky RL, Mendelsohn J, et al. Impact of precision medicine in diverse cancers: a meta-analysis of phase II clinical trials. J Clin Oncol. 2015;33(32):3817–25. https://doi.org/10.1200/JCO.2015.61.5997.
14. Dine J, Gordon R, Shames Y, Kasler MK, Barton-Burke M. Immune checkpoint inhibitors: an innovation in immunotherapy for the treatment and management of patients with cancer. Asia-Pacific Journal of Oncology Nursing. 2017;4(2):127–35. https://doi.org/10.4103/apjon.apjon_4_17.
15. Shih K, Arkenau H-T, Infante JR. Clinical impact of checkpoint inhibitors as novel cancer therapies. Drugs. 2014;74(17):1993–2013. https://doi.org/10.1007/s40265-014-0305-6.
16. Melero I, Gaudernack G, Gerritsen W, Huber C, Parmiani G, Scholl S, et al. Therapeutic vaccines for cancer: an overview of clinical trials. Nat Rev Clin Oncol. 2014;11(9):509–24. https://doi.org/10.1038/nrclinonc.2014.111.
17. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer. 2012;12(4):252–64. https://doi.org/10.1038/nrc3239.
18. Postow MA, Callahan MK, Wolchok JD. Immune checkpoint blockade in cancer therapy. J Clin Oncol. 2015;33(17):1974–82. https://doi.org/10.1200/JCO.2014.59.4358.
19. Sharma P, Allison JP. Immune checkpoint targeting in cancer therapy: toward combination strategies with curative potential. Cell. 2015;161(2):205–14. https://doi.org/10.1016/j.cell.2015.03.030.
20. Mina M, et al. Conditional selection of genomic alterations dictates cancer evolution and oncogenic dependencies. *Cancer Cell*. **32**:155–168.e156.
21. Austin C, Kusumoto F. The application of big data in medicine: current implications and future directions. J Int Cardiac Electrophysiol. 2016;47(1):51–9. https://doi.org/10.1007/s10840-016-0104-y.
22. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The cancer imaging archive (TCIA): maintaining and operating a public information repository. J Digit Imaging. 2013;26(6):1045–57. https://doi.org/10.1007/s10278-013-9622-7.
23. Kellis M, Wold B, Snyder MP, Bernstein BE, Kundaje A, Marinov GK, et al. Defining functional DNA elements in the human genome. Proc Natl Acad Sci. 2014;111(17):6131–8. https://doi.org/10.1073/pnas.1318948111.
24. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The genotype-tissue expression (GTEx) project. Nat Gen. 2013;45(6):580
25. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, et al. The NIH roadmap epigenomics mapping consortium. Nat Biotechnol. 2010;28(10):1045–8. https://doi.org/10.1038/nbt1010-1045.
26. Weinstein JN, et al. The cancer genome atlas pan-cancer analysis project. Nat Genet. 2013;45(10):1113–20. https://doi.org/10.1038/ng.2764.
27. Li J, Lu Y, Akbani R, Ju Z, Roebuck PL, Liu W, et al. TCPA: a resource for cancer functional proteomics data. Nat Meth. 2013;10(11):1046–7. https://doi.org/10.1038/nmeth.2650.
28. Shoemaker RH. The NCI60 human tumour cell line anticancer drug screen. Nat Rev Cancer. 2006;6(10):813–23. https://doi.org/10.1038/nrc1951.
29. Horgan RP, Kenny LC. 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics. The Obstetrician & Gynaecologist. 2011;13(3):189–95. https://doi.org/10.1576/toag.13.3.189.27672.
30. Yang Y, Dong X, Xie B, Ding N, Chen J, Li Y, et al. Databases and web tools for cancer genomics study. Genomics, Proteomics & Bioinformatics. 2015;13(1):46–50. https://doi.org/10.1016/j.gpb.2015.01.005.
31. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012;483(7391):603–7. https://doi.org/10.1038/nature11003.
32. Garnett MJ, Edelman EJ, Heidorn SJ, Greenman CD, Dastur A, Lau KW, et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature. 2012;483(7391):570–5. https://doi.org/10.1038/nature11005.
33. Firmino M, Angelo G, Morais H, Dantas MR, Valentim R. Computer-aided detection (CADe) and diagnosis (CADx) system for lung cancer with likelihood of malignancy. Biomed Eng Online. 2016;15(1):2. https://doi.org/10.1186/s12938-015-0120-7.
34. 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. https://doi.org/10.1016/j.mri.2012.06.010.
35. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2016;278(2):563–77. https://doi.org/10.1148/radiol.2015151169.
36. Wong AJ, Kanwar A, Mohamed AS, Fuller CD. Radiomics in head and neck cancer: from exploration to application. Trans Cancer Res. 2016;5(4):371–82. https://doi.org/10.21037/tcr.2016.07.18.

37. Zhang, Y. Image processing. (Walter de Gruyter GmbH & Co KG, 2017).
38. Friedman J, Hastie T, Tibshirani R. The elements of statistical learning, vol. 1: Springer series in statistics, New York; 2001.
39. Giraud C. Introduction to high-dimensional statistics. Vol. 138. CRC Press; 2014.
40. Murdoch TB, Detsky AS. The inevitable application of big data to health care. JAMA. 2013;309(13):1351–2. https://doi.org/10.1001/jama.2013.393.
41. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. J Am Med Inform Assoc. 2013;20(1):144–51. https://doi.org/10.1136/amiajnl-2011-000681.
42. Ross MK, Wei W, Ohno-Machado L. "Big data" and the electronic health record. Yearb Med Inform. 2014;9(1):97–104. https://doi.org/10.15265/IY-2014-0003.
43. Denny JC. Chapter 13: mining electronic health records in the genomics era. PLoS Comput Biol. 2012;8(12):e1002823. https://doi.org/10.1371/journal.pcbi.1002823.
44. Jpc Rodrigues J, de la Torre I, Fernández G, López-Coronado M. Analysis of the security and privacy requirements of cloud-based electronic health records systems. J Med Internet Res. 2013;15(8):e186. https://doi.org/10.2196/jmir.2494.
45. Schmitt P, Mandel J, Guedj M. A comparison of six methods for missing data imputation. J Biomet Biostat. 2015;6:1.
46. Manning CD and Schütze, H. Foundations of statistical natural language processing, Vol. 999. MIT Press; 1999.
47. Poncelet, P. Data mining patterns: new methods and applications. IGI Global; 2007.
48. Goldstein BA, Navar AM, Pencina MJ, Ioannidis JPA. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. J Am Med Inform Assoc. 2017;24(1):198–208. https://doi.org/10.1093/jamia/ocw042.
49. Mandrekar SJ, Grothey A, Goetz MP, Sargent DJ. Clinical trial designs for prospective validation of biomarkers. Am J Pharmacogenomics. 2005;5(5):317–25. https://doi.org/10.2165/00129785-200505050-00004.
50. Saber H, Gudi R, Manning M, Wearne E, Leighton JK. An FDA oncology analysis of immune activating products and first-in-human dose selection. Regul Toxicol Pharmacol. 2016;81:448–56. https://doi.org/10.1016/j.yrtph.2016.10.002.
51. Peer D, et al. Nanocarriers as an emerging platform for cancer therapy. Nat Nano. 2007;2:751–60.
52. DeMets DL, Ellenberg SS. Data monitoring committees—expect the unexpected. N Engl J Med. 2016;375(14):1365–71. https://doi.org/10.1056/NEJMra1510066.
53. Chow S-C and Liu J-P. Design and analysis of clinical trials: concepts and methodologies. Vol. 507. John Wiley & Sons; 2008.
54. Treweek S, Altman DG, Bower P, Campbell M, Chalmers I, Cotton S, et al. Making randomised trials more efficient: report of the first meeting to discuss the Trial Forge platform. Trials. 2015;16(1):261. https://doi.org/10.1186/s13063-015-0776-0.
55. Porta, M. A dictionary of epidemiology. Oxford university press; 2008.
56. Jasen P. Breast cancer and the politics of abortion in the United States. Med Hist. 2005;49(04):423–44. https://doi.org/10.1017/S0025727300009145.
57. Collett D. Modelling survival data in medical research. CRC press; 2015.
58. Hosmer Jr, D.W., Lemeshow, S. & Sturdivant, R.X. Applied logistic regression. Vol. 398. John Wiley & Sons; 2013.
59. Gaber MM, Zaslavsky A, Krishnaswamy S. Mining data streams: a review. SIGMOD Rec. 2005;34(2):18–26. https://doi.org/10.1145/1083784.1083789.
60. Gama, J. & Gaber, M.M. Learning from data streams: processing techniques in sensor networks. Springer; 2007.
61. Modi PK, Farber NJ, Singer EA. Precision oncology: identifying predictive biomarkers for the treatment of metastatic renal cell carcinoma. Translational cancer research. 2016;5(S1):S76–80. https://doi.org/10.21037/tcr.2016.06.05.
62. Nair S, Liew C, Khor TO, Cai L, Kong A-N. Differential signaling regulatory networks governing hormone refractory prostate cancers. J Chin Pharm Sci. 2014;23:511–24.
63. Vogelstein B, Kinzler K. Cancer genes and the pathways they control. Nat Med. 2004;10(8):789–99. https://doi.org/10.1038/nm1087.
64. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. Nature. 2013;501(7467):338–45. https://doi.org/10.1038/nature12625.
65. Nair S, Iyer A, Vijay V, Bandlamudi S, Llerena A. Pharmacokinetics and systems pharmacology of monoclonal antibody olaratumab for inoperable soft tissue sarcoma. Advances in Modern Oncology Research. 2017;3:114–25.
66. Sun X, Hu B. Mathematical modeling and computational prediction of cancer drug resistance. Brief Bioinform. 2017; https://doi.org/10.1093/bib/bbx065.
67. Kerr KM, Tsao MS, Nicholson AG, Yatabe Y, Wistuba II, Hirsch FR, et al. Programmed death-ligand 1 immunohistochemistry in lung cancer: in what state is this art? J Thorac Oncol. 2015;10(7):985–9. https://doi.org/10.1097/JTO.0000000000000526.
68. Patel SP, Kurzrock R. PD-L1 expression as a predictive biomarker in cancer immunotherapy. Mol Cancer Ther. 2015;14(4):847–56. https://doi.org/10.1158/1535-7163.MCT-14-0983.
69. Johnson DB, Peng C, Sosman JA. Nivolumab in melanoma: latest evidence and clinical potential. Ther Adv Med Oncol. 2015;7(2):97–106. https://doi.org/10.1177/1758834014567469.
70. Nair S. Pharmacometrics and systems pharmacology of immune checkpoint inhibitor nivolumab in cancer translational medicine. Adv Mod Oncol Res. 2016;2(1):18–31. https://doi.org/10.18282/amor.v2.i1.46.
71. Nalejska E, Maczynska E, Lewandowska MA. Prognostic and predictive biomarkers: tools in personalized oncology. Mol Diag Ther. 2014;18(3):273–84. https://doi.org/10.1007/s40291-013-0077-9.
72. Yates LR, Campbell PJ. Evolution of the cancer genome. Nat Rev Genet. 2012;13(11):795–806. https://doi.org/10.1038/nrg3317.
73. Tourneau CL, et al. The spectrum of clinical trials aiming at personalizing medicine. Chin Clin Oncol. 2014;3
74. Wassmer,G and Brannath, W. Group sequential and confirmatory adaptive designs in clinical trials. Springer; 2016.
75. Jennison C and Turnbull BW. Group sequential methods with applications to clinical trials. CRC Press; 1999.
76. Simon N, Simon R. Adaptive enrichment designs for clinical trials. Biostatistics (Oxford, England). 2013;14:613–25.
77. Hommel G. Adaptive modifications of hypotheses after an interim analysis. Biom J. 2001;43(5):581–9. https://doi.org/10.1002/1521-4036(200109)43:5<581::AID-BIMJ581>3.0.CO;2-J.
78. Krisam J, Kieser M. Decision rules for subgroup selection based on a predictive biomarker. J Biopharm Stat. 2014;24(1):188–202. https://doi.org/10.1080/10543406.2013.856018.
79. Bauer P, Kohne K. Evaluation of experiments with adaptive interim analyses. Biometrics. 1994;50(4):1029–41. https://doi.org/10.2307/2533441.
80. Brannath W, Gutjahr G, Bauer P. Probabilistic foundation of confirmatory adaptive designs. J Am Stat Assoc. 2012;107(498):824–32. https://doi.org/10.1080/01621459.2012.682540.
81. Ji Z, Yan K, Li W, Hu H, Zhu X. Mathematical and computational modeling in complex biological systems. Biomed Res Int. 2017;2017:16.

82. Li Z, Li P, Krishnan A, Liu J. Large-scale dynamic gene regulatory network inference combining differential equation models with local dynamic Bayesian network analysis. Bioinformatics. 2011;27(19):2686–91. https://doi.org/10.1093/bioinformatics/btr454.
83. Spencer SL, Berryman MJ, García JA, Abbott D. An ordinary differential equation model for the multistep transformation to cancer. J Theor Biol. 2004;231(4):515–24. https://doi.org/10.1016/j.jtbi.2004.07.006.
84. Riemann, R.-C. Modelling of concurrent systems: structural and semantical methods in the high level Petri net calculus. Herbert Utz Verlag; 1999.
85. Nikolaev A and Vázquez-Abad FJ. In: Winter Simulation Conference (WSC), 2015 1471–1482 (IEEE, 2015).
86. Esparza J, Nielsen M. Decidability issues for Petri nets. BRICS Report Series. 1994;1(8) https://doi.org/10.7146/brics.v1i8.21662.
87. Huang S. Genomics, complexity and drug discovery: insights from Boolean network models of cellular regulation. Pharmacogenomics. 2001;2(3):203–22. https://doi.org/10.1517/14622416.2.3.203.
88. Albert R, Othmer HG. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in Drosophila melanogaster. J Theor Biol. 2003;223(1):1–18. https://doi.org/10.1016/S0022-5193(03)00035-3.
89. Li J, Bench AJ, Vassiliou GS, Fourouclas N, Ferguson-Smith AC, Green AR. Imprinting of the human L3MBTL gene, a polycomb family member located in a region of chromosome 20 deleted in human myeloid malignancies. Proc Natl Acad Sci U S A. 2004;101(19):7341–6. https://doi.org/10.1073/pnas.0308195101.
90. Mahesh S, Mangla N, Pooja V, Bhyratae S. Reconstruction of Gene Regulatory Network to identify prognostic molecular markers of the reactive stroma of breast and prostate cancer using information theoretic approach. Int J Innovative Res Comput Commun Eng. 2015;3:303–10.
91. Barnes, D.J. & Chu, D. Introduction to modeling for biosciences. (Springer Science & Business Media, 2010).
92. Michor F, Beal K. Improving cancer treatment via mathematical modeling: surmounting the challenges is worth the effort. Cell. 2015;163(5):1059–63. https://doi.org/10.1016/j.cell.2015.11.002.
93. Dréau, D., Stanimirov, D., Carmichael, T. & Hadzikadic, M. In: Bioinformatics and Computational Biology: First International Conference, BICoB 2009, New Orleans, LA, USA, April 8–10, 2009. Proceedings. S. Rajasekaran (ed.), 187–198 (Springer Berlin Heidelberg, Berlin, Heidelberg; 2009).
94. Cox, D.R. & Oakes, D. Analysis of survival data. Vol. 21. CRC Press; 1984).
95. Gray, R.J. Taylor L & Francis, 2002.
96. Zhang X, Li Y, Akinyemiju T, Ojesina AI, Buckhaults P, Liu N, et al. Pathway-structured predictive model for cancer survival prediction: a two-stage approach. Genetics. 2017;205(1):89–100. https://doi.org/10.1534/genetics.116.189191.
97. Stein LD. The case for cloud computing in genome informatics. Genome Biol. 2010;11(5):207. https://doi.org/10.1186/gb-2010-11-5-207.
98. Freimuth RR, Formea CM, Hoffman JM, Matey E, Peterson JF, Boyce RD. Implementing genomic clinical decision support for drug-based precision medicine. CPT. 2017;6(3):153–5. https://doi.org/10.1002/psp4.12173.
99. Nedungadi P, Jayakumar A, Raman R. Personalized health monitoring system for managing well-being in rural areas. J Med Syst. 2018;42(1):22. https://doi.org/10.1007/s10916-017-0854-9.
100. van Nimwegen KJ, et al. Is the $1000 genome as near as we think? A cost analysis of next-generation sequencing. Clin Chem. 2016;62(11):1458–64. https://doi.org/10.1373/clinchem.2016.258632.
101. Martz CA, et al. Systematic identification of signaling pathways with potential to confer anticancer drug resistance. *Sci Signal*. 2014;7:ra121.
102. Au TH, Wang K, Stenehjem D, Garrido-Laguna I. Personalized and precision medicine: integrating genomics into treatment decisions in gastrointestinal malignancies. J Gastrointest Oncol. 2017;8(3):387–404. https://doi.org/10.21037/jgo.2017.01.04.
103. Weng C. Optimizing clinical research participant selection with informatics. Trends Pharmacol Sci. 2015;36(11):706–9. https://doi.org/10.1016/j.tips.2015.08.007.
104. Azuaje F. What does systems biology mean for biomarker discovery? Exp Opin Med Diagn. 2010;4(1):1–10. https://doi.org/10.1517/17530050903468709.
105. Edelman EJ, Guinney J, Chi J-T, Febbo PG, Mukherjee S. Modeling cancer progression via pathway dependencies. PLoS Comput Biol. 2008;4(2):e28. https://doi.org/10.1371/journal.pcbi.0040028.