

REVIEW

Open Access



# Are innovation and new technologies in precision medicine paving a new era in patients centric care?

Attila A. Seyhan<sup>1,2\*</sup> and Claudio Carini<sup>3</sup>

## Abstract

Healthcare is undergoing a transformation, and it is imperative to leverage new technologies to generate new data and support the advent of precision medicine (PM). Recent scientific breakthroughs and technological advancements have improved our understanding of disease pathogenesis and changed the way we diagnose and treat disease leading to more precise, predictable and powerful health care that is customized for the individual patient. Genetic, genomics, and epigenetic alterations appear to be contributing to different diseases. Deep clinical phenotyping, combined with advanced molecular phenotypic profiling, enables the construction of causal network models in which a genomic region is proposed to influence the levels of transcripts, proteins, and metabolites. Phenotypic analysis bears great importance to elucidate the pathophysiology of networks at the molecular and cellular level. Digital biomarkers (BMs) can have several applications beyond clinical trials in diagnostics—to identify patients affected by a disease or to guide treatment. Digital BMs present a big opportunity to measure clinical endpoints in a remote, objective and unbiased manner. However, the use of “omics” technologies and large sample sizes have generated massive amounts of data sets, and their analyses have become a major bottleneck requiring sophisticated computational and statistical methods. With the wealth of information for different diseases and its link to intrinsic biology, the challenge is now to turn the multi-parametric taxonomic classification of a disease into better clinical decision-making by more precisely defining a disease. As a result, the big data revolution has provided an opportunity to apply artificial intelligence (AI) and machine learning algorithms to this vast data set. The advancements in digital health opportunities have also arisen numerous questions and concerns on the future of healthcare practices in particular with what regards the reliability of AI diagnostic tools, the impact on clinical practice and vulnerability of algorithms. AI, machine learning algorithms, computational biology, and digital BMs will offer an opportunity to translate new data into actionable information thus, allowing earlier diagnosis and precise treatment options. A better understanding and cohesiveness of the different components of the knowledge network is a must to fully exploit the potential of it.

**Keywords:** Precision medicine, Personalized medicine, Biomarkers, Modeling and simulation, Digital biomarkers, Artificial intelligence, Machine learning, Deep phenotyping, Cancer, Immuno-oncology, Autoimmune and inflammatory diseases, Diabetes, Genetics, Epigenetics, Genomics, Transcriptomics, microRNAs, miRNAs, Proteomics

\*Correspondence: Attila\_Seyhan@brown.edu; Attila\_seyhan@yahoo.com

<sup>1</sup> Department of Pathology and Laboratory Medicine, Division of Biology and Medicine, Brown University, Providence, RI 02903, USA

Full list of author information is available at the end of the article



## Introduction

Today, the practice of medicine remains largely empirical; physicians generally rely on patterns matching to establish a diagnosis based on a combination of the patients' medical history, physical examination, and laboratory data. Thus, a given treatment is often based on physicians past experience with similar patients. One consequence of this is that a blockbuster gets prescribed for a "typical patient" with a specific disease. According to this paradigm, treatment decision is driven by trial and error and the patient occasionally becomes the victim of unpredictable side effects, or poor or no efficacy for a drug that theoretically works in some people affected by that specific disease.

Greater use of BMs [1, 2] and companion diagnostics (CDX) can now enable a shift from empirical medicine to precision medicine (PM) (the right medicine, for the right patient, at the right dose, at the right time). It is conceivable that, in the immediate future, physicians will be moving away from the concept of "one size fits all" and shift instead to PM.

It is generally known that the response of a specific treatment varies across the heterogeneity of a population with good and poor responders. Patients and treatment response differ because of variables like genetic predisposition, heterogeneity of the cohorts, ethnicity, slow vs. fast metabolizers, epigenetic factors, early vs. late stage of the disease. These parameters have an effect on whether a given individual will be a good or poor responder to a specific treatment.

The goal of PM is to enable clinicians to quickly, efficiently and accurately predict the most appropriate course of action for a patient. To achieve this, clinicians are in need of tools that are both compatible with their clinical workflow and economically feasible. Those tools can simplify the process of managing the biological complexity that underlies human diseases. To support the creation and refinement of those tools, a PM ecosystem is in continuous development and is the solution to the problem. The PM ecosystem is beginning to link and share information among clinicians, laboratories, research enterprises, and clinical-information-system developers. It is expected that these efforts will create the foundation of a continuously evolving health-care system that is capable of significantly accelerating the advancement of PM technologies.

Precision medicine highlights the importance of coupling established clinical indexes with molecular profiling in order to craft diagnostic, prognostic and therapeutic strategies specific for the needs of each group of patients. A correct interpretation of the data is a must for the best use of the PM ecosystem. The PM ecosystem combines omics and clinical data to

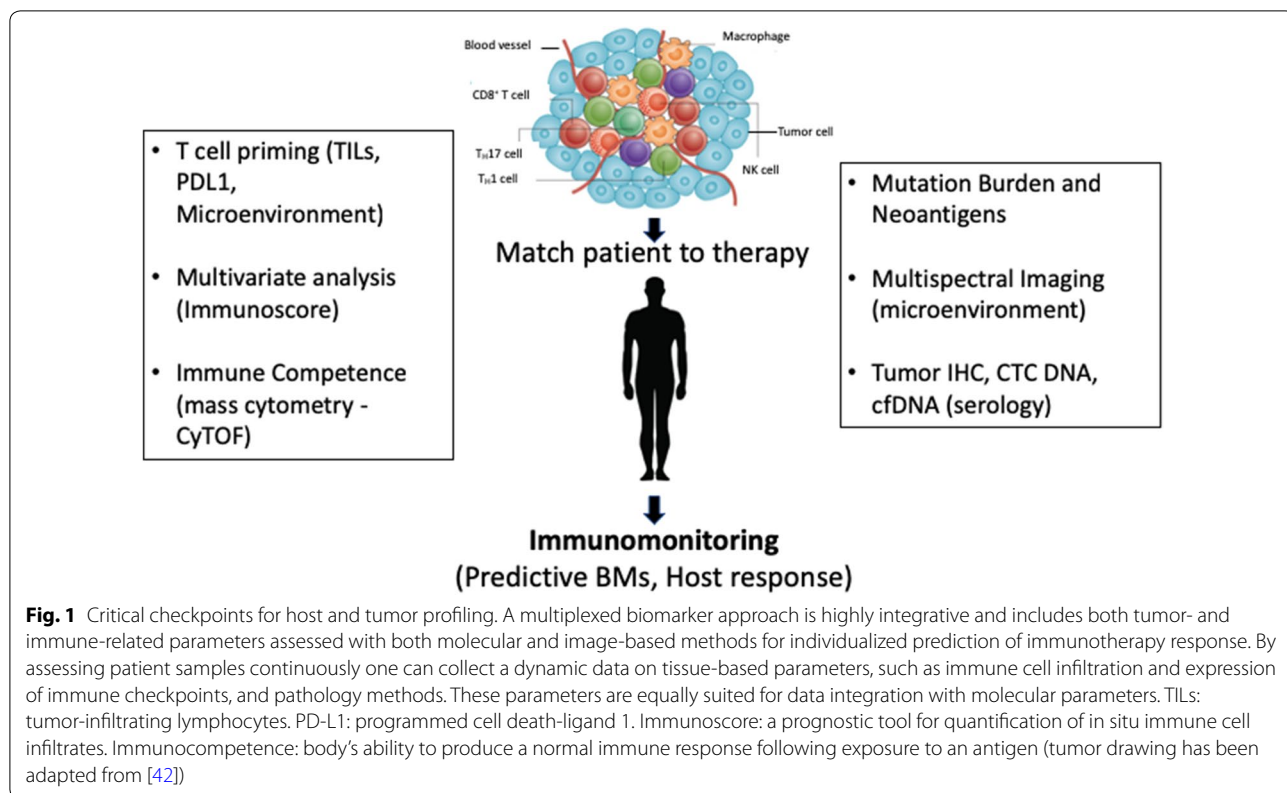
determine the best course of action to be taken for each specific patient group.

Currently, a drug gets approved after a lengthy regulatory process. One way to address this problem, is to focus on selected group of patients thus, Phase III clinical studies can be conducted with a small group of patients rather than thousands and thousands of patients typically needed for the Phase III studies. This approach should potentially guarantee a more rapid and expeditious way to perform drug development of next-generation pharmacotherapy. A narrower focus on a specific patient's group at the stage of the regulatory approval process should facilitate streamlining the regulatory approval resulting in a greater clinical and economic success.

The shift towards a deeper understanding of disease based on molecular biology will also inevitably lead to a new, more precise disease's classification, incorporating new molecular knowledge to generate a new taxonomy. This change will result in a revised classification of intrinsic biology, leading to revisions of diseases signs and symptoms. For this change to occur, however, larger data bases, accessible to all, will be needed that dynamically incorporate new information.

The emerging use of personalized laboratory medicine makes use of a multitude of testing options that can more precisely pinpoint management needs of individual groups of patients. PM seeks to dichotomize patient populations in those who might benefit from a specific treatment (responders) and those for whom a benefit is improbable (non-responders). Defining cut-off points and criteria for such a dichotomy is difficult. Treatment recommendations are often generated using algorithms based on individual somatic genotype alterations. However, tumors often harbor multiple drivers' mutations (owing to intra- and inter-tumoral heterogeneity). Physicians, therefore, need to combine different streams of evidence to prioritize their choice of treatment. The implementation of PM often relies on a fragmented landscape of evidences making hard for physicians to select among different diagnostic tools and treatment options.

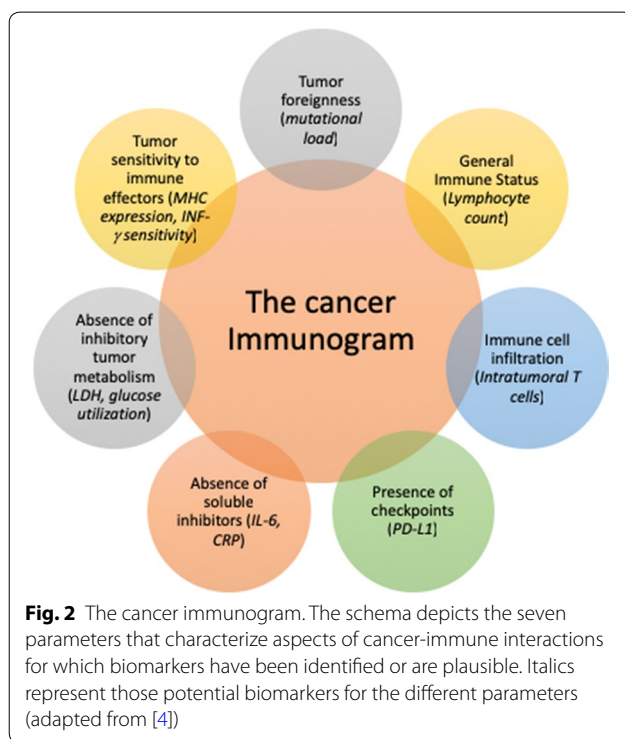
In the case of cancer immunotherapy, predictive biomarkers (BM) for immunotherapy differ from the traditional BM used for targeted therapies. The complexity of the tumor microenvironment (TME), the immune response and molecular profiling requires a more holistic approach than the use of a single analyte BM [3]. To cope with this challenge, researchers have adopted multiplexing approach, where multiple BMs are used to empower more accurate patient stratification [4]. To select specific patient's groups for immunotherapy, histological analysis now include concomitant analysis of immuno-oncology BMs, such as PD-L1 and immune cell infiltrates (Fig. 1) as well as more comprehensive immune and tumor-related pathways



(the “Cancer Immunogram”) (Fig. 2) [4, 5]. In the case of cancer immunotherapy, multiplexed immunoprofiling generating a comprehensive biomarker dataset that can correlated with clinical parameters is key for the success of PM.

**Patient stratification for precision medicine**

In traditional drug development, patients with a disease are enrolled randomly to avoid bias, using an “all comers” approach with the assumption that the enrolled patients are virtually homogeneous. The reason for random enrollment is to ensure a wide representation of the general population. In reality, we never perform clinical trials for randomly selected patients, but rather we apply various types of enrichments to patients’ enrolment by applying specific inclusion and exclusion criteria. Despite all of those efforts to increase the enrichment, the population that ultimately gets selected for the study can be rather heterogeneous with respect to drug-metabolizing capabilities, environmental conditions (e.g. diet, smoking habit, lifestyle etc.), or previous exposure to medication(s) as well as individuals genetic and epigenetic make-up. By using BMs to better characterize molecular, genetic, and epigenetic makeup of patients, drug developers have been trying to establish a more objective approach.



The use of patient stratification is to separate probable responders from non-responders. A prospective stratification can result in a smaller and shorter clinical study compared to those needed for randomly selected patients.

Minimally, stratification can speed up approval for drug candidates intended for a subset of patients, while leaving the door open for further tests and market expansion in the more heterogeneous population of patients. Maximally, it can unmask a useful therapeutic agent that otherwise would be lost in the noise generated by the non-responders, as was the case for instance of trastuzumab and gefitinib [6].

Thus, clinical trials could be shorter, given a quicker determination on the efficacy of the new molecular entity. Today, the major focus of research is to identify the molecular causes of differential therapeutic responses across patient populations. It is now clear that patients affected by a disease show significant response heterogeneity to a given treatment. Advances in understanding the mechanisms underlying diseases and drug response are increasingly creating opportunities to match patients with therapies that are more likely to be efficacious and safer.

Furthermore, patient stratification has a considerable economic impact on the model of the pharmaceutical industry. By identifying the populations likely to benefit from a new therapy, drug development costs will be reduced and the risk of treating non-responders will be minimized. Advances in “omics” technologies (e.g. epigenomics, genomics, transcriptomics, proteomics, metabolomics, etc.), also called, systems-based approach [7], are now utilized to identify molecular targets including BMs [1, 2] that can reveal the disease state or the ability to respond to a specific treatment, thus providing scientists and clinicians to generate a learning dataset consisting of molecular insights of the disease pathogenesis.

A search of the relevant literature will reveal an abundance of publications related to BMs [8]. However, as previously reported by Poste in 2011 [9] more than 150,000 articles have described thousands of BMs however, only approximately 100 BMs are routinely used in the clinical practice. As to date, over 355 new non-traditional BMs (i.e. pharmacogenomic BM-drug pairs) have been described in drug labels ([www.fda.gov/drugs/scienceresearch/ucm572698.htm](http://www.fda.gov/drugs/scienceresearch/ucm572698.htm)). Table 1 lists 355 pharmacogenomic BMs as of Dec. 2018, linked to drugs with pharmacogenomic information found in the drug labeling (Drugs@FDA; <https://www.fda.gov/drugs/scienceresearch/ucm572698.htm>). Those BMs include germline or somatic gene variants (i.e. polymorphisms, mutations), functional deficiencies with a genetic etiology, altered gene expression signatures, and chromosomal

abnormalities. The list also includes selected protein BMs that are used to select treatments for specific patient's groups.

Moreover, as reported recently by Burke [10] there are more than 768,000 papers indexed in PubMed.gov directly related to BMs (<https://www.amplion.com/biomarker-trends/biomarker-panels-the-good-the-bad-and-the-ugly/>).

All the data collected so far have shown insufficient linkages between BMs and disease pathogenesis resulting in the failure of many BMs as well as drug targets. It is critical to link the target to the disease pathogenesis thus, enabling the development of better and more precise therapies by pre-selecting responders to treatment.

### **Biomarkers and decision making**

BMs have been used to improve patient's stratification and/or develop targeted therapies facilitating the decision-making process throughout the new drug development process. BMs constitute a rational approach which, at its most optimal, reflects both the biology of the disease and the effectiveness of the drug candidate. Also, adding the appropriate BMs to a drug-development strategy enables the concept of ‘fail fast, fail early’; thus, allowing early identification of the high proportion of compounds that fail during drug development. Reducing human exposure to drugs with low efficacy or safety concerns allows to shift resources to drugs that have a higher chance of success. Identification of BMs helpful for a quick go-no-go decision early in the drug development process is critical for enhancing the probability of success of a drug.

Traditionally, clinical trial end-points, such as morbidity and mortality, often require extended timeframes and may be difficult to evaluate. Imaging-based BMs are providing objective end-points that may be confidently evaluated in a reasonable timeframe. However, imaging techniques are rather expensive and often very impractical especially in specific geographical area.

Despite all of these, BMs are essential for deciding which patients should receive a specific treatment. Table 1 illustrates a number of pharmacogenomic BMs in drug labeling. As of December 2018, approximately 355 pharmacogenomic BMs are linked to drugs with pharmacogenomic information found in the drug labeling. These BMs include germline or somatic gene variants (i.e. polymorphisms, mutations), functional deficiencies with a genetic etiology, altered gene expression signatures, and chromosomal abnormalities, and selected protein BMs that are used to select treatments for patients.

Pre-clinical BMs are essential, as long they translate into clinical markers. Which is often is not the case. Several reasons can be offered to explain why many clinical

**Table 1 Pharmacogenomic BMs in drug labeling**

Drug	Therapeutic area	Biomarker	Labeling sections
Abacavir	Infectious Diseases	HLA-B	Boxed warning, dosage and administration, contraindications, warnings and precautions
Abemaciclib (1)	Oncology	ESR (Hormone receptor)	Indications and usage, adverse reactions, clinical studies
Abemaciclib (2)	Oncology	ERBB2 (HER2)	Indications and usage, adverse reactions, clinical studies
Ado-Trastuzumab Emtansine	Oncology	ERBB2 (HER2)	Indications and usage, warnings and precautions, adverse reactions, clinical pharmacology, clinical studies
Afatinib	Oncology	EGFR	Indications and usage, dosage and administration, adverse reactions, clinical studies
Alectinib	Oncology	ALK	Indications and usage, dosage and administration, adverse reactions, clinical pharmacology, clinical studies
Amitriptyline	Psychiatry	CYP2D6	Precautions
Anastrozole	Oncology	ESR, PGR (Hormone receptor)	Indications and usage, adverse reactions, drug interactions, clinical studies
Arformoterol (1)	Pulmonary	UGT1A1	Clinical pharmacology
Arformoterol (2)	Pulmonary	CYP2D6	Clinical pharmacology
Aripiprazole	Psychiatry	CYP2D6	Dosage and administration, use in specific populations, clinical pharmacology
Aripiprazole lauroxil	Psychiatry	CYP2D6	Dosage and administration, use in specific populations, clinical pharmacology
Arsenic trioxide	Oncology	PML-RARA	Indications and usage
Atezolizumab	Oncology	CD274 (PD-L1)	Indications and usage, adverse reactions, clinical pharmacology, clinical studies
Atomoxetine	Psychiatry	CYP2D6	Dosage and administration, warnings and precautions, adverse reactions, drug interactions, clinical pharmacology
Ascorbic acid, PEG-3350, potassium chloride, sodium ascorbate, sodium chloride, and sodium sulfate	Gastroenterology	G6PD	Warnings and precautions
Avatrombopag (1)	Hematology	F2 (Prothrombin)	Warnings and precautions
Avatrombopag (2)	Hematology	F5 (Factor V Leiden)	Warnings and precautions
Avatrombopag (3)	Hematology	PROC	Warnings and precautions
Avatrombopag (4)	Hematology	PROS1	Warnings and precautions

**Table 1 (continued)**

<b>Drug</b>	<b>Therapeutic area</b>	<b>Biomarker</b>	<b>Labeling sections</b>
Avatrombopag (5)	Hematology	SERPINC1 (Antithrombin III)	Warnings and precautions
Avelumab	Oncology	CD274 (PD-L1)	Clinical studies
Azathioprine	Rheumatology	TPMT	Dosage and administration, warnings, precautions, drug interactions, adverse reactions, clinical pharmacology
Belinostat	Oncology	UGT1A1	Dosage and administration, clinical pharmacology
Binimetinib (1)	Oncology	BRAF	Indications and usage, dosage and administration, warnings and precautions, adverse reactions, use in specific populations, clinical studies
Binimetinib (2)	Oncology	UGT1A1	Clinical pharmacology
Blinatumomab	Oncology	BCR-ABL1 (Philadelphia chromosome)	Indications and usage, clinical studies
Boceprevir	Infectious Diseases	IFNL3 (IL28B)	Clinical pharmacology
Bosutinib	Oncology	BCR-ABL1 (Philadelphia chromosome)	Indications and usage, adverse reactions, use in specific populations, clinical studies
Brentuximab vedotin	Oncology	ALK	Clinical studies
Brexiprazole	Psychiatry	CYP2D6	Dosage and administration, use in specific populations, clinical pharmacology
Brigatinib	Oncology	ALK	Indications and usage, adverse reactions, clinical studies
Brivaracetam	Neurology	CYP2C19	Clinical pharmacology
Busulfan	Oncology	BCR-ABL1 (Philadelphia chromosome)	Clinical studies
Cabozantinib	Oncology	RET	Clinical studies
Capecitabine	Oncology	DPYD	Warnings and precautions, patient counseling information
Carbamazepine (1)	Neurology	HLA-B	Boxed warning, warnings, precautions
Carbamazepine (2)	Neurology	HLA-A	Warnings
Carlumic Acid	Inborn errors of metabolism	NAGS	Indications and usage, warnings and precautions, use in specific populations, clinical pharmacology, clinical studies
Cariprazine	Psychiatry	CYP2D6	Clinical pharmacology
Carisoprodol	Rheumatology	CYP2C19	Use in specific populations, clinical pharmacology
Carvedilol	Cardiology	CYP2D6	Drug interactions, clinical pharmacology

**Table 1 (continued)**

<b>Drug</b>	<b>Therapeutic area</b>	<b>Biomarker</b>	<b>Labeling sections</b>
Celecoxib	Rheumatology	CYP2C9	Dosage and administration, use in specific populations, clinical pharmacology
Ceritinib	Oncology	ALK	Indications and usage, dosage and administration, adverse reactions, clinical studies
Cerliponase alfa	Inborn errors of metabolism	TPP1	Indications and usage, use in specific populations, clinical studies
Cetuximab (1)	Oncology	EGFR	Indications and usage, dosage and administration, warnings and precautions, adverse reactions, clinical studies
Cetuximab (2)	Oncology	RAS	Indications and usage, dosage and administration, warnings and precautions, adverse reactions, clinical studies
Cevimeline	Dental	CYP2D6	Precautions
Chloroquine	Infectious diseases	G6PD	Precautions
Chlorpropamide	Endocrinology	G6PD	Precautions
Cisplatin	Oncology	TPMT	Adverse reactions
Citalopram (1)	Psychiatry	CYP2C19	Dosage and administration, warnings, clinical pharmacology
Citalopram (2)	Psychiatry	CYP2D6	Clinical pharmacology
Clobazam	Neurology	CYP2C19	Dosage and administration, use in specific populations, clinical pharmacology
Clomipramine	Psychiatry	CYP2D6	Precautions
Clopidogrel	Cardiology	CYP2C19	Boxed warning, warnings and precautions, clinical pharmacology
Clozapine	Psychiatry	CYP2D6	Dosage and administration, use in specific populations, clinical pharmacology
Cobimetinib	Oncology	BRAF	Indications and usage, dosage and administration, adverse reactions, clinical studies
Codeine	Anesthesiology	CYP2D6	Boxed warning, warnings and precautions, use in specific populations, patient counseling information
Crizotinib (1)	Oncology	ALK	Indications and usage, dosage and administration, adverse reactions, use in specific populations, clinical pharmacology, clinical studies
Crizotinib (2)	Oncology	ROS1	Indications and usage, dosage and administration, adverse reactions, use in specific populations, clinical studies

**Table 1 (continued)**

Drug	Therapeutic area	Biomarker	Labeling sections
Dabrafenib (1)	Oncology	BRAF	Indications and usage, dosage and administration, warnings and precautions, adverse reactions, clinical pharmacology, clinical studies, patient counseling information
Dabrafenib (2)	Oncology	G6PD	Warnings and precautions, adverse reactions, patient counseling information
Dabrafenib (3)	Oncology	RAS	Dosage and administration, warnings and precautions
Daclatasvir	Infectious diseases	IFNL3 (IL28B)	Clinical studies
Dapsone (1)	Dermatology	G6PD	Warnings and precautions, use in specific populations
Dapsone (2)	Dermatology	Nonspecific (Congenital Methemoglobinemia)	Warnings and precautions
Dapsone (3)	Infectious Diseases	G6PD	Precautions, adverse reactions, overdose
Darifenacin	Urology	CYP2D6	Clinical pharmacology
Dasabuvir, ombitasvir, paritaprevir, and ritonavir	Infectious Diseases	IFNL3 (IL28B)	Clinical studies
Dasatinib	Oncology	BCR-ABL1 (Philadelphia chromosome)	Indications and usage, dosage and administration, warnings and precautions, adverse reactions, clinical studies
Denileukin diftitox	Oncology	IL2RA (CD25 antigen)	Indications and usage, warnings and precautions, clinical studies
Desipramine	Psychiatry	CYP2D6	Precautions
Desflurane	Anesthesiology	Nonspecific (Genetic Susceptibility to Malignant Hyperthermia)	Contraindications
Desvenlafaxine	Psychiatry	CYP2D6	Clinical pharmacology
Deutetrabenazine	Neurology	CYP2D6	Dosage and administration, warnings and precautions, use in specific populations, clinical pharmacology
Dexlansoprazole	Gastroenterology	CYP2C19	Drug interactions, clinical pharmacology
Dextromethorphan and quinidine	Neurology	CYP2D6	Warnings and precautions, clinical pharmacology
Diazepam	Neurology	CYP2C19	Clinical pharmacology
Dinutuximab	Oncology	MYCN	Clinical studies
Dolutegravir	Infectious Diseases	UGT1A1	Clinical pharmacology
Doxepin (1)	Psychiatry	CYP2D6	Clinical pharmacology
Doxepin (2)	Psychiatry	CYP2C19	Clinical pharmacology
Dronabinol	Gastroenterology	CYP2C9	Use in specific populations, clinical pharmacology
Drospirenone and ethinyl estradiol	Gynecology	CYP2C19	Clinical pharmacology
Duloxetine	Psychiatry	CYP2D6	Drug interactions



**Table 1 (continued)**

<b>Drug</b>	<b>Therapeutic area</b>	<b>Biomarker</b>	<b>Labeling sections</b>
Durvalumab	Oncology	CD274 (PD-L1)	Clinical pharmacology, clinical studies
Efavirenz	Infectious diseases	CYP2B6	Clinical pharmacology
Elbasvir and grazoprevir	Infectious diseases	IFNL3 (IL28B)	Clinical studies
Eliglustat	Inborn errors of metabolism	CYP2D6	Indications and usage, dosage and administration, contraindications, warnings and precautions, drug interactions, use in specific populations, clinical pharmacology, clinical studies
Elosulfase	Inborn errors of metabolism	GALNS	Indications and usage, warnings and precautions, use in specific populations, clinical pharmacology, clinical studies
Eltrombopag (1)	Hematology	F5 (Factor V Leiden)	Warnings and precautions
Eltrombopag (2)	Hematology	SERPINC1 (Antithrombin III)	Warnings and precautions
Enasidenib	Oncology	IDH2	Indications and usage, dosage and administration, clinical pharmacology, clinical studies
Encorafenib	Oncology	BRAF	Indications and usage, dosage and administration, warnings and precautions, adverse reactions, use in specific populations, clinical pharmacology, clinical studies
Enflurane	Anesthesiology	Nonspecific (genetic susceptibility to malignant hyperthermia)	Contraindications
Erlotinib	Oncology	EGFR	Indications and usage, dosage and administration, adverse reactions, clinical studies
Erythromycin and sulfisoxazole	Infectious Diseases	G6PD	Precautions
Escitalopram (1)	Psychiatry	CYP2D6	Drug interactions
Escitalopram (2)	Psychiatry	CYP2C19	Adverse reactions
Esomeprazole	Gastroenterology	CYP2C19	Drug interactions, clinical pharmacology
Eteplirsen	Neurology	DMD	Indications and usage, adverse reactions, use in specific populations, clinical studies
Everolimus (1)	Oncology	ERBB2 (HER2)	Indications and usage, dosage and administration, warnings and precautions, adverse reactions, drug interactions, use in specific populations, clinical studies
Everolimus (2)	Oncology	ESR (hormone receptor)	Indications and usage, dosage and administration, warnings and precautions, adverse reactions, drug interactions, use in specific populations, clinical studies

**Table 1 (continued)**

<b>Drug</b>	<b>Therapeutic area</b>	<b>Biomarker</b>	<b>Labeling sections</b>
Exemestane	Oncology	ESR, PGR (hormone receptor)	Indications and usage, dosage and administration, clinical studies
Fesoterodine	Urology	CYP2D6	Drug interactions, clinical pharmacology
Filbanerin (1)	Gynecology	CYP2C9	Clinical pharmacology
Filbanerin (2)	Gynecology	CYP2C19	Adverse reactions, use in specific populations, clinical pharmacology
Filbanerin (3)	Gynecology	CYP2D6	Clinical pharmacology
Fluorouracil (1)	Dermatology	DPYD	Contraindications, warnings
Fluorouracil (2)	Oncology	DPYD	Warnings and precautions, patient counseling information
Fluoxetine	Psychiatry	CYP2D6	Precautions, clinical pharmacology
Flurbiprofen	Rheumatology	CYP2C9	Clinical pharmacology
Fluvoxamine	Psychiatry	CYP2D6	Drug interactions
Formoterol (1)	Pulmonary	CYP2D6	Clinical pharmacology
Formoterol (2)	Pulmonary	CYP2C19	Clinical pharmacology
Fulvestrant (1)	Oncology	ERBB2 (HER2)	Indications and usage, adverse reactions, clinical studies
Fulvestrant (2)	Oncology	ESR, PGR (Hormone Receptor)	Indications and usage, adverse reactions, clinical pharmacology, clinical studies
Galantamine	Neurology	CYP2D6	Clinical pharmacology
Gefitinib (1)	Oncology	EGFR	Indications and usage, dosage and administration, clinical studies
Gefitinib (2)	Oncology	CYP2D6	Clinical pharmacology
Glimepiride	Endocrinology	G6PD	Warnings and precautions, adverse reactions
Glipizide	Endocrinology	G6PD	Precautions
Glyburide	Endocrinology	G6PD	Precautions
Hydralazine	Cardiology	Nonspecific (NAT)	Clinical pharmacology
Ibrutinib (1)	Oncology	Chromosome 17p	Indications and usage, clinical studies
Ibrutinib (2)	Oncology	Chromosome 11q	Clinical studies
Iloperidone	Psychiatry	CYP2D6	Dosage and administration, warnings and precautions, drug interactions, clinical pharmacology
Imatinib (1)	Oncology	KIT	Indications and usage, dosage and administration, clinical studies

**Table 1 (continued)**

Drug	Therapeutic area	Biomarker	Labeling sections
Imatinib (2)	Oncology	BCR-ABL1 (Philadelphia chromosome)	Indications and usage, dosage and administration, warnings and precautions, adverse reactions, use in specific populations, clinical pharmacology, clinical studies
Imatinib (3)	Oncology	PDGFRB	Indications and usage, dosage and administration, clinical studies
Imatinib (4)	Oncology	FIP1L1-PDGFR	Indications and usage, dosage and administration, clinical studies
Imipramine	Psychiatry	CYP2D6	Precautions
Indacaterol	Pulmonary	UGT1A1	Clinical pharmacology
Inotuzumab Ozogamicin	Oncology	BCR-ABL1 (Philadelphia chromosome)	Clinical studies
Irinotecan	Oncology	UGT1A1	Dosage and administration, warnings and precautions, clinical pharmacology
Isoflurane	Anesthesiology	Nonspecific (genetic susceptibility to malignant hyperthermia)	Contraindications
Isoniazid, pyrazinamide, and rifampin	Infectious Diseases	Nonspecific (NAT)	Clinical pharmacology
Isosorbide dinitrate	Cardiology	CYP5R	Overdosage
Isosorbide mononitrate	Cardiology	CYP5R	Overdosage
Ivacaftor	Pulmonary	CFTR	Indications and usage, adverse reactions, use in specific populations, clinical pharmacology, clinical studies
Ivacaftor and lumacaftor	Pulmonary	CFTR	Indications and usage, adverse reactions, use in specific populations, clinical studies
Ivacaftor and tezacaftor	Pulmonary	CFTR	Indications and usage, adverse reactions, use in specific populations, clinical pharmacology, clinical studies
Lacosamide	Neurology	CYP2C19	Clinical pharmacology
Lansoprazole	Gastroenterology	CYP2C19	Drug interactions, clinical pharmacology
Lapatinib (1)	Oncology	ERBB2 (HER2)	Indications and usage, dosage and administration, adverse reactions, use in specific populations, clinical studies
Lapatinib (2)	Oncology	ESR, PGR (Hormone Receptor)	Indications and usage, dosage and administration, adverse reactions, use in specific populations, clinical studies
Lapatinib (3)	Oncology	HLA-DQA1, HLA-DRB1	Clinical pharmacology
Ledipasvir and sofosbuvir	Infectious diseases	IFNL3 (IL28B)	Clinical studies

**Table 1 (continued)**

<b>Drug</b>	<b>Therapeutic area</b>	<b>Biomarker</b>	<b>Labeling sections</b>
Lenalidomide	Hematology	Chromosome 5q	Boxed warning, indications and usage, adverse reactions, use in specific populations, clinical studies
Lesinurad	Rheumatology	CYP2C9	Drug interactions, clinical pharmacology
Letrozole	Oncology	ESR, PGR (Hormone Receptor)	Indications and usage, adverse reactions, clinical studies
Lidocaine and prilocaine (1)	Anesthesiology	Nonspecific (congenital methemoglobinemia)	Warnings and precautions
Lidocaine and prilocaine (2)	Anesthesiology	G6PD	Warnings and precautions, clinical pharmacology
Lofexidine	Anesthesiology	CYP2D6	Use in specific populations
Mafenide	Infectious diseases	G6PD	Warnings, adverse reactions
Meclizine	Neurology	CYP2D6	Clinical pharmacology
Mercaptopurine (1)	Oncology	TPMT	Dosage and administration, warnings and precautions, adverse reactions, clinical pharmacology
Mercaptopurine (2)	Oncology	NUDT15	Dosage and administration, warnings and precautions, clinical pharmacology
Methylene blue	Hematology	G6PD	Contraindications, warnings and precautions
Metoclopramide (1)	Gastroenterology	CYB5R	Precautions, overdose
Metoclopramide (2)	Gastroenterology	G6PD	Precautions, overdose
Metoprolol	Cardiology	CYP2D6	Drug interactions, clinical pharmacology
Midostaurin (1)	Oncology	FLT3	Indications and usage, dosage and administration, adverse reactions, clinical studies
Midostaurin (2)	Oncology	NPM1	Clinical studies
Midostaurin (3)	Oncology	KIT	Clinical studies
Mirabegron	Urology	CYP2D6	Clinical pharmacology
Modafinil	Psychiatry	CYP2D6	Clinical pharmacology
Mycophenolic acid	Transplantation	HPRT1	Warnings and precautions
Nalidixic acid	Infectious diseases	G6PD	Precautions, adverse reactions
Nebivolol	Cardiology	CYP2D6	Dosage and administration, clinical pharmacology
Nefazodone	Psychiatry	CYP2D6	Precautions
Neratinib (1)	Oncology	ERBB2 (HER2)	Indications and usage, adverse reactions, clinical studies
Neratinib (2)	Oncology	ESR, PGR (Hormone Receptor)	Clinical studies
Nilotinib (1)	Oncology	BCR-ABL1 (Philadelphia chromosome)	Indications and usage, dosage and administration, warnings and precautions, adverse reactions, use in specific populations, clinical studies
Nilotinib (2)	Oncology	UGT1A1	Clinical pharmacology

**Table 1 (continued)**

<b>Drug</b>	<b>Therapeutic area</b>	<b>Biomarker</b>	<b>Labeling sections</b>
Niraparib	Oncology	BRCA	Clinical studies
Nitrofurantoin	Infectious Diseases	G6PD	Warnings, adverse reactions
Nivolumab (1)	Oncology	BRAF	Indications and usage, adverse reactions, clinical studies
Nivolumab (2)	Oncology	CD274 (PD-L1)	Clinical pharmacology, clinical studies
Nivolumab (3)	Oncology	Microsatellite instability, mismatch repair	Indications and usage, use in specific populations, clinical pharmacology, clinical studies
Nortriptyline	Psychiatry	CYP2D6	Precautions
Obinutuzumab	Oncology	MS4A1 (CD20 antigen)	Clinical studies
Olaparib	Oncology	BRCA	Indications and usage, dosage and administration, warnings and precautions, adverse reactions, clinical studies
Olaratumab	Oncology	PDGFRA	Clinical studies
Omacetaxine	Oncology	BCR-ABL1 (Philadelphia chromosome)	Clinical studies
Ombitasvir, paritaprevir, and ritonavir	Infectious diseases	IFNL3 (IL28B)	Clinical studies
Omeprazole	Gastroenterology	CYP2C19	Drug interactions, clinical pharmacology
Ondansetron	Gastroenterology	CYP2D6	Clinical pharmacology
Osimertinib	Oncology	EGFR	Indications and usage, dosage and administration, adverse reactions, clinical studies
Oxcarbazepine	Neurology	HLA-B	Warnings and precautions
Palbociclib (1)	Oncology	ESR (Hormone Receptor)	Indications and usage, adverse reactions, clinical studies
Palbociclib (2)	Oncology	ERBB2 (HER2)	Indications and usage, adverse reactions, clinical studies
Palonosetron	Gastroenterology	CYP2D6	Clinical pharmacology
Panitumumab (1)	Oncology	EGFR	Adverse reactions, clinical pharmacology, clinical studies
Panitumumab (2)	Oncology	RAS	Indications and usage, dosage and administration, warnings and precautions, adverse reactions, clinical studies
Pantoprazole	Gastroenterology	CYP2C19	Clinical pharmacology
Parathyroid hormone	Inborn errors of metabolism	CASR	Indications and usage, clinical studies
Paroxetine	Psychiatry	CYP2D6	Drug interactions
Pazopanib (1)	Oncology	UGT1A1	Clinical pharmacology

**Table 1 (continued)**

<b>Drug</b>	<b>Therapeutic area</b>	<b>Biomarker</b>	<b>Labeling sections</b>
Pazopanib (2)	Oncology	HLA-B	Clinical pharmacology
Peginterferon alfa-2b	Infectious Diseases	IFNL3 (IL28B)	Clinical pharmacology
Pegloticase	Rheumatology	G6PD	Boxed warning, contraindications, warnings and precautions, patient counseling information
Pembrolizumab (1)	Oncology	BRAF	Adverse reactions, clinical studies
Pembrolizumab (2)	Oncology	CD274 (PD-L1)	Indications and usage, dosage and administration, use in specific populations, clinical studies
Pembrolizumab (3)	Oncology	Microsatellite Instability, Mismatch Repair	Indications and usage, Dosage and administration, use in specific populations, clinical studies
Perphenazine	Psychiatry	CYP2D6	Precautions, clinical pharmacology
Pertuzumab (1)	Oncology	ERBB2 (HER2)	Indications and usage, warnings and precautions, adverse reactions, clinical pharmacology, clinical studies
Pertuzumab (2)	Oncology	ESR, PGR (Hormone Receptor)	Clinical studies
Phenytoin (1)	Neurology	CYP2C9	Clinical pharmacology
Phenytoin (2)	Neurology	CYP2C19	Clinical pharmacology
Phenytoin (3)	Neurology	HLA-B	Warnings
Pimozide	Psychiatry	CYP2D6	Dosage and administration, precautions
Piroxicam	Rheumatology	CYP2C9	Clinical pharmacology
Ponatinib	Oncology	BCR-ABL1 (Philadelphia chromosome)	Indications and usage, warnings and precautions, adverse reactions, use in specific populations, clinical studies
Prasugrel (1)	Cardiology	CYP2C19	Use in specific populations, clinical pharmacology, clinical studies
Prasugrel (2)	Cardiology	CYP2C9	Use in specific populations, clinical pharmacology, clinical studies
Prasugrel (3)	Cardiology	CYP3A5	Use in specific populations, clinical pharmacology, clinical studies
Prasugrel (4)	Cardiology	CYP2B6	Use in specific populations, clinical pharmacology, clinical studies
Primaquine (1)	Infectious Diseases	G6PD	Contraindications, warnings, precautions, adverse reactions, overdose
Primaquine (2)	Infectious Diseases	CYB5R	Precautions, adverse reactions
Propafenone	Cardiology	CYP2D6	Dosage and administration, warnings and precautions, drug interactions, clinical pharmacology

**Table 1 (continued)**

<b>Drug</b>	<b>Therapeutic area</b>	<b>Biomarker</b>	<b>Labeling sections</b>
Propranolol	Cardiology	CYP2D6	Clinical pharmacology
Protriptyline	Psychiatry	CYP2D6	Precautions
Quinidine	Cardiology	CYP2D6	Precautions
Quinine sulfate (1)	Infectious Diseases	G6PD	Contraindications
Quinine sulfate (2)	Infectious Diseases	CYP2D6	Drug interactions
Rabeprazole	Gastroenterology	CYP2C19	Clinical pharmacology
Raltegravir	Infectious Diseases	UGT1A1	Clinical pharmacology
Rasburicase (1)	Oncology	G6PD	Boxed warning, contraindications, warnings and precautions
Rasburicase (2)	Oncology	CYB5R	Boxed warning, contraindications, warnings and precautions
Ribociclib (1)	Oncology	ESR, PGR (Hormone Receptor)	Indications and usage, clinical studies
Ribociclib (2)	Oncology	ERBB2 (HER2)	Indications and usage, clinical studies
Risperidone	Psychiatry	CYP2D6	Drug interactions, clinical pharmacology
Rituximab	Oncology	MS4A1 (CD20 antigen)	Indications and usage, dosage and administration, adverse reactions, use in specific populations, clinical studies
Rosuvastatin	Endocrinology	SLCO1B1	Clinical pharmacology
Rucaparib (1)	Oncology	BRCA	Indications and usage, dosage and administration, adverse reactions, use in specific populations, clinical studies
Rucaparib (2)	Oncology	CYP2D6	Clinical pharmacology
Rucaparib (3)	Oncology	CYP1A2	Clinical pharmacology
Sevoflurane	Anesthesiology	Nonspecific (Genetic Susceptibility to Malignant Hyperthermia)	Warnings
Simeprevir	Infectious Diseases	IFNL3 (IL28B)	Clinical pharmacology, clinical studies
Sodium nitrite	Toxicology	G6PD	Warnings and precautions
Sofosbuvir	Infectious Diseases	IFNL3 (IL28B)	Clinical studies
Sofosbuvir and velpatasvir	Infectious Diseases	IFNL3 (IL28B)	Clinical studies
Sofosbuvir, velpatasvir, and voxilaprevir	Infectious Diseases	IFNL3 (IL28B)	Clinical studies
Succimer	Hematology	G6PD	Clinical pharmacology

**Table 1 (continued)**

<b>Drug</b>	<b>Therapeutic area</b>	<b>Biomarker</b>	<b>Labeling sections</b>
Succinylcholine	Anesthesiology	BCHE	Warnings, precautions
Sulfamethoxazole and trimethoprim (1)	Infectious Diseases	G6PD	Precautions
Sulfamethoxazole and trimethoprim (2)	Infectious Diseases	Nonspecific (NAT)	Precautions
Sulfasalazine (1)	Gastroenterology	G6PD	Precautions
Sulfasalazine (2)	Gastroenterology	Nonspecific (NAT)	Clinical pharmacology
Tamoxifen (1)	Oncology	ESR, PGR (Hormone receptor)	Indications and usage, precautions, adverse reactions, clinical studies
Tamoxifen (2)	Oncology	F5 (Factor V Leiden)	Warnings
Tamoxifen (3)	Oncology	F2 (Prothrombin)	Warnings
Telaprevir	Infectious Diseases	IFNL3 (IL28B)	Clinical pharmacology, clinical studies
Tetrabenazine	Neurology	CYP2D6	Dosage and administration, warnings and precautions, use in specific populations, clinical pharmacology
Thioguanine (1)	Oncology	TPMT	Dosage and administration, warnings, precautions, clinical pharmacology
Thioguanine (2)	Oncology	NUDT15	Dosage and administration, warnings, precautions, clinical pharmacology
Thioridazine	Psychiatry	CYP2D6	Contraindications, warnings, precautions
Ticagrelor	Cardiology	CYP2C19	Clinical pharmacology
Tolterodine	Urology	CYP2D6	Precautions, clinical pharmacology
Tramadol	Anesthesiology	CYP2D6	Boxed warning, warnings, precautions, use in specific populations, clinical pharmacology
Trametinib (1)	Oncology	BRAF	Indications and usage, dosage and administration, adverse reactions, clinical pharmacology, clinical studies, patient counseling information
Trametinib (2)	Oncology	G6PD	Adverse reactions
Trametinib (3)	Oncology	RAS	Warnings and precautions
Trastuzumab (1)	Oncology	ERBB2 (HER2)	Indications and usage, warnings and precautions, clinical pharmacology, clinical studies
Trastuzumab (2)	Oncology	ESR, PGR (Hormone receptor)	Clinical studies
Tretinoin	Oncology	PML-RARA	Indications and usage, warnings, clinical pharmacology
Trimipramine	Psychiatry	CYP2D6	Precautions



**Table 1 (continued)**

Drug	Therapeutic area	Biomarker	Labeling sections
Umeclidinium	Pulmonary	CYP2D6	Clinical pharmacology
Ustekinumab	Dermatology and gastroenterology	IL12A, IL12B, IL23A	Warnings and precautions
Valbenazine	Neurology	CYP2D6	Dosage and administration, warnings and precautions, use in specific populations, clinical pharmacology
Valproic acid (1)	Neurology	POLG	Boxed warning, contraindications, warnings and precautions
Valproic acid (2)	Neurology	Nonspecific (urea cycle disorders)	Contraindications, warnings and precautions
Vemurafenib (1)	Oncology	BRAF	Indications and usage, dosage and administration, warnings and precautions, adverse reactions, use in specific populations, clinical pharmacology, clinical studies, patient counseling information
Vemurafenib (2)	Oncology	RAS	Warnings and precautions, adverse reactions
Venlafaxine	Psychiatry	CYP2D6	Precautions
Venetoclax	Oncology	Chromosome 17p	Indications and usage, dosage and administration, use in specific populations, clinical studies
Voriconazole	Infectious Diseases	CYP2C19	Clinical pharmacology
Vortioxetine	Psychiatry	CYP2D6	Dosage and administration, clinical pharmacology
Warfarin (1)	Hematology	CYP2C9	Dosage and administration, drug interactions, clinical pharmacology
Warfarin (2)	Hematology	VKORC1	Dosage and administration, clinical pharmacology
Warfarin (3)	Hematology	PROS1	Warnings and precautions
Warfarin (4)	Hematology	PROC	Warnings and precautions

The table lists 355 BMs as of December 2018, linked to drugs with pharmacogenomic information found in the drug labeling. The list of BMs includes but are not limited to germline or somatic gene variants (i.e. polymorphisms, mutations), functional deficiencies with a genetic etiology, altered gene expression signatures, and chromosomal abnormalities, and selected protein BMs that are used to select treatments for patients (adapted from Drugs@FDA; <https://www.fda.gov/drugs/science/research/ucm572698.htm>)

studies have failed to identify BMs ability to predict treatment efficacy or disease modification including lack of statistical power, lack of validation standards [11] and pharmacogenetic heterogeneity of patient groups [12].

#### **Genomics, epigenetics, and microRNAs as emerging biomarkers in cancer, diabetes, autoimmune and inflammatory diseases**

Biomarkers with the potential to identify early stages of disease for example pre-neoplastic disease or very early stages of cancer are of great promise to improve patient survival. The concept of liquid biopsy refers to a minimally invasive collection and analysis of molecules that can be isolated from body fluids, primarily whole blood, serum, plasma, urine and saliva, and others. A myriad of circulating molecules such as cell-free DNA (cf-DNA), cell-free RNA (cf-RNA) including microRNAs (miRNAs), circulating tumor cells (CTC), circulating tumor proteins, and extracellular vesicles, more specifically exosomes, have been explored as biomarkers [13].

Genetic and epigenetic alterations including DNA methylation and altered miRNA expression might be contributing to several autoimmune diseases, cancer, transplantation, and infectious diseases. For example in a recent study in rheumatoid arthritis (RA), de la Rica et al. [14] has identified epigenetic factors involved in RA, and hence conducted DNA methylation and miRNA expression profiling of a set of RA synovial fibroblasts and compared the results with those obtained from osteoarthritis (OA) patients with a normal phenotype. In this study, researchers identified changes in novel key genes including IL6R, CAPN8, and DPP4, as well as several HOX genes. Notably, many genes modified by DNA methylation were inversely correlated with expression miRNAs. A comprehensive analysis revealed several miRNAs that are controlled by DNA methylation, and genes that are regulated by DNA methylation and targeted by miRNAs were of potential use as clinical markers. The study found that several genes including Stat4 and TRAF1-C5 were identified as risk factors contributing to RA and other autoimmune diseases such as SLE [15, 16]. RA is also strongly associated with the inherited tissue type MHC antigen HLA-DR4 and the genes PTPN22 and PAD14 [15]. DNA methylation screening identified genes undergoing DNA methylation-mediated silencing including IL6R, CAPN8 and DPP4, as well as several HOX genes; and a panel of miRNAs that are controlled by DNA methylation, and genes that are regulated by DNA methylation and are targeted by miRNAs.

Likewise, changes in miRNA levels in blood and other body fluids (miRNAs) have been linked to a variety of autoimmune diseases [17] including: (i) Type 1 diabetes, miR-342, miR-191, miR-375 and miR-21 and miR-510

and others [18–20]; (ii) Type 2 diabetes, miR-30, miR-34a, miR-145 and miR-29c, miR-138, -192, -195, -320b, and let-7a, (iii) prediabetes (miR-7, miR-152 and miR-192) [21, 22] and insulin resistance (miR-24, miR-30d, miR-146a), obesity and metabolic diseases [19–26] (iv) Multiple sclerosis (MS), miR-326 [27], miR-17-5p [28]; (v) Rheumatoid Arthritis (RA), miR-146a, miR-155 and miR-16 [29, 30]; (vi) Primary biliary cirrhosis, miR-122a, miR-26a, miR-328, miR-299-5p [31]; (vii) Sjögren's syndrome, miR-17-92 [17]; (viii) SLE, miR-146a [32], miR-516-5p, miR-637 [33]; and (ix) Psoriasis, miR-203, miR-146a, miR125b, miR21 [34].

In the case of RA, alterations in several miRNAs expression patterns including miR-146a, miRNA-155, miRNA-124a, miR-203, miR-223, miR-346, miR-132, miR-363, miR-498, miR-15a, and miR-16 were documented in several tissue samples of RA patients. The polymorphisms present in these miRNAs and their targets have also been associated with RA or other autoimmune diseases [19, 35]. Several reports have shown altered miRNA expression in the synovium of patients with RA [36]. For example, elevated expression of miR-346 was found in Lipopolysaccharide activated RA fibroblast-like synoviocytes (FLS) [37]. Moreover, miR-124 was found at lower levels in RA FLS in comparison with FLS from patients with OA [38]. miR-146a has been found to be elevated in human RA synovial tissue and its expression is induced by the pro-inflammatory cytokines i.e. tumor necrosis factor and interleukin1 $\beta$  [29]. Furthermore, miR-146, miR-155, and miR-16 were all elevated in the peripheral blood of RA patients with the active disease rather than inactive disease [30] suggesting that these miRNAs may serve as potential disease activity markers.

The epigenetic regulation of DNA processes has been extensively studied over the past 15 years in cancer, where DNA methylation and histone modification, nucleosome remodeling and RNA mediated targeting regulate many biological processes that are crucial to the genesis of cancer. The first evidence indicating of an epigenetic link with cancer were studied derived from DNA methylation. Though many of the initial studies were purely correlative, however, they did highlight a potential connection between epigenetic pathways and cancer. These preliminary results were confirmed by recent results from the International Cancer Genome Consortium (ICGC).

Compilation of the epigenetic regulators mutated in cancer highlights histone acetylation and methylation as the most widely affected epigenetic pathways. Deep sequencing technologies aimed at mapping chromatin modifications have begun to shed some lights on the origin of epigenetic abnormalities in cancer. Several pieces of evidence are now highlighting that dysregulation of the epigenetic pathways can lead to cancer. All the evidence

collected thus far along with clinical and preclinical results observed with epigenetic drugs against chromatin regulators, point to the necessity of embracing a central role of epigenetics in cancer. Unfortunately, those studies are far too many to be comprehensively described in this review.

Furthermore, cancer cell lines have been used to identify potential novel biomarkers for drug resistance and novel targets and pathways for drug repurposing. For example, previously we conducted a functional shRNA screen combined with a lethal dose of neratinib to discover chemo-resistant interactions with neratinib. We identified a collection of genes whose inhibition by RNAi led to neratinib resistance including genes involved in oncogenesis, transcription factors, cellular ion transport, protein ubiquitination, cell cycle, and genes known to interact with breast cancer-associated genes [39]. These novel mediators of cellular resistance to neratinib could lead to their use as patient or treatment selection biomarkers.

In addition, we undertook a genome-wide pooled lentiviral shRNA screen to identify synthetic lethal or enhancer (synthetic modulator screen) genes that interact with sub-effective doses of neratinib in a human breast cancer cell line. We discovered a diverse set of genes whose depletion selectively impaired or enhanced the viability of cancer cells in the presence of neratinib. Further examination of these genes and pathways led to a rationale for the treatment of cells with either paclitaxel or cytarabine in combination with neratinib which resulted in a strong antiproliferative effect. Notably, our findings support a paclitaxel and neratinib phase II clinical trial in breast cancer patients [40].

### Biomarker multiplexing

Multiple biomarkers are used to empower more accurate patient stratification. To improve patient stratification for immunotherapy, the analysis of immuno-oncology biomarkers, like PD-L1, as well as a more comprehensive analysis of the immune and tumor-related pathways (the “Cancer Immunogram”) (Fig. 2) [4] has to be used for a better patient stratification in future immunotherapy trials [5]. The “Cancer Immunogram” includes tumor foreignness, immune status, immune cell infiltration, absence of checkpoints, absence of soluble inhibitors, absence of inhibitory tumor metabolism, and tumor sensitivity to immune effectors as the most important predictors of immunotherapy response in a single tissue sample [5]. As depicted in Fig. 2, The “Cancer Immunogram” integrates both tumor- and immune-related characteristics assessed with both molecular and image-based methods for individualized prediction of immunotherapy response. By evaluating dynamic data on tissue-based

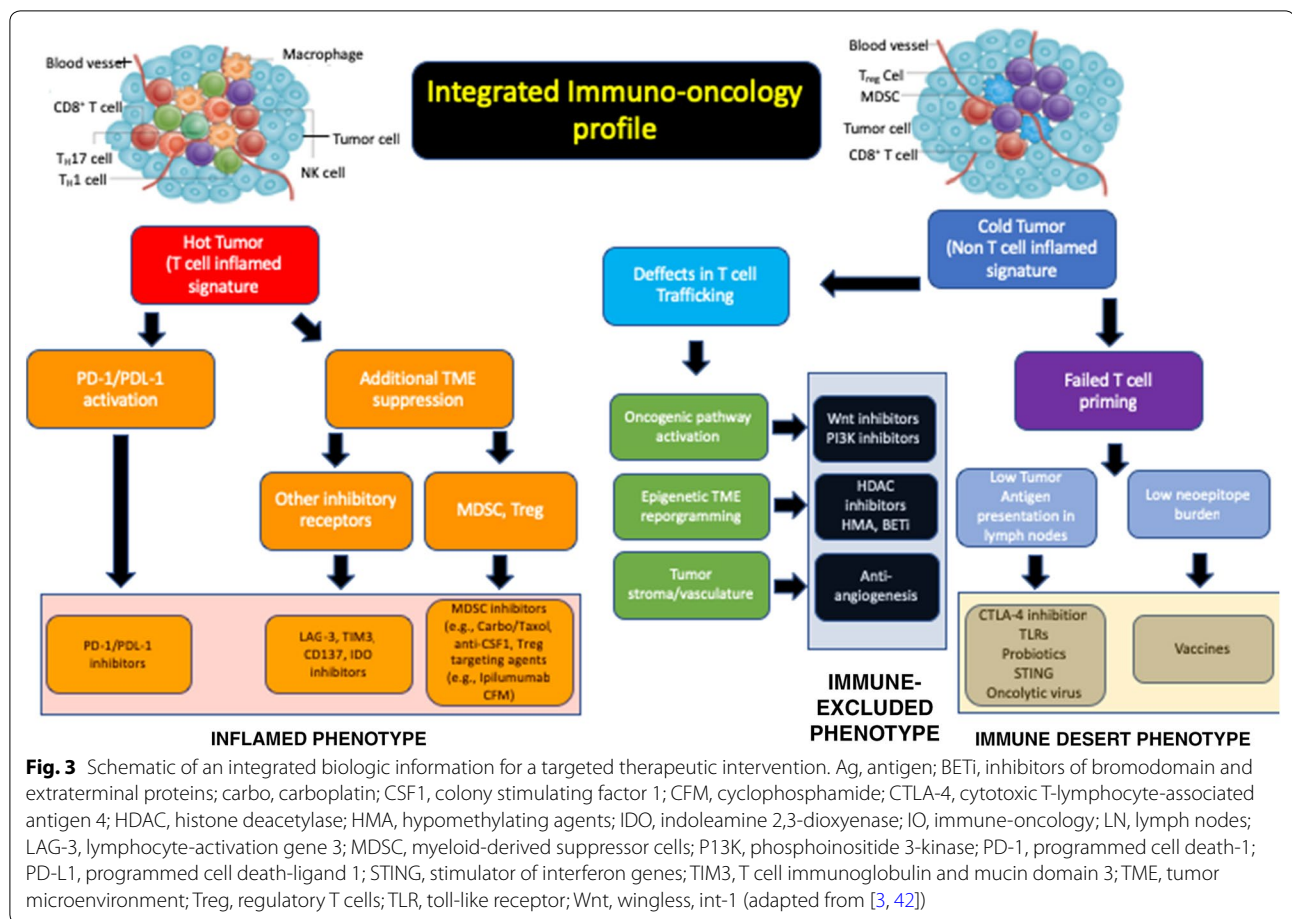
parameters, (e.g., immune cell infiltration and expression of immune checkpoints), quantitative pathology methods are ideally suited for data integration with molecular parameters.

As illustrated in Fig. 3, and reported in a recent article [3], the utility of this approach to organize and integrate the biologic information into a useful and informative single assay able to inform and influence drug development, personalized therapy strategy and selection of specific patient populations. The authors [3] suggest that anti-cancer immunity can be histologically segregated into three main phenotypes: (1) the inflamed phenotype (“hot” tumors); (2) the immune-excluded phenotype (“cold” tumors); and (3) the immune-desert phenotype (“cold” tumors) [41, 42] (Fig. 3). Each tumor phenotype is associated with specific underlying biological and pathological mechanisms that may determine the success of the host immune response and immunotherapy or other therapeutic modalities to fight cancer. Identifying these mechanisms at the level of the individual groups of patients and selecting those patients with similar tumor phenotype is critical for the selection of specific patient populations both for the development as well as implementation of therapeutic interventions.

### Digital biomarkers

Digital BMs are defined as an objective, quantifiable physiological and behavioral data that are collected and measured by means of digital devices. The data collected is typically used to explain, influence and/or predict health-related outcomes. Increasingly, many smartphone apps are also available for health management with or without connection to these sensor devices [43, 44]. There are approx. 300,000 health apps and 340+ (CK personal communication) sensor devices available today and the number of apps is doubling every 2 years. Recently, a new class of wearable smartphone-coupled devices such as smart watches have been widely available. These devices offer new, and more practical opportunities not without limitations [44]. As those wearable devices and their corresponding apps continue to develop and evolve, there will be a need for a more dedicated research and digital expert assessment to evaluate different healthcare applications as well as assess the limitations and the risks of impinging on the individual privacy and data safety.

This surge in technology has made it possible for ‘consumers’ to track their health but also represents an interesting opportunity to monitor healthcare and clinical trials. Data collected about a patient’s activity and vital signs can be used to get an idea about the patient’s health status and disease progression on a daily basis. However, the problem is that a majority of these apps and devices



are meant for wellness purposes and not intended to diagnose or treat diseases.

As reported previously in the literature [5], and shown Figs. 1 and 2, recent advances in electronic data collection will be instrumental in our ability to digitize and process large collections of tissue slides and molecular diagnostic profiling. The evolving field of machine learning and artificial intelligence with the support of human interpretation will have a dramatic impact on the field [45, 46].

This field has already generated tangible results. Indeed, medical device companies (e.g., Philips, GE, and Leica) are developing new imaging technologies for digital pathology to detect digital biomarkers, while a number of Information Technology (IT) companies (e.g., Google, IBM, and Microsoft, or PathAI) are developing tools, such as machine learning and artificial intelligence (AI) for big data analysis and integrated decision making.

Pharmaceutical companies are also moving in the same direction. For example, FDA clearance for the VENTANA MMR IHC Panel for patients diagnosed with colorectal cancer (CRC) developed by Roche is a demonstration of

these efforts [5]. Thus, developing digital biomarkers, big data analysis and interpretation will be beneficial in the new era of PM.

**How can wearable help in clinical trials and healthcare?**

In a typical clinical trial or in a clinical setting, the patient visits the hospital not more than once per month or less. So, the clinician can observe the signs and symptoms of the patient only during this visit and has almost no visibility on how the patient is doing for the majority of the time outside the clinic. If digital BMs are used, the patient can perform these tests using smartphones or sensors in the comfort of his/her home. For example, in a Parkinson's disease trial various aspects of the patient's health can be captured in a remote study using smartphone-based apps. This allows the collection of quantitative and unbiased data on a frequent or almost continuous basis. The clinician can get almost real-time feedback on each patient, whether they are getting better or worse. This feedback can help to inform the study protocol or even halt the study if the drug doesn't seem to be working on most of the patients.

The Clinical Trials Transformation Initiative (CTTI) provides a framework and detailed guidance for developing digital BMs. They also outline the benefits of using digital BMs in clinical trials such as being patient-centric while also making faster decisions that save time and costs.

#### ***Develop and validate digital biomarkers***

The first and most important consideration in developing digital BMs is not which device to use, but rather deciding which disease symptoms to capture that best represent the disease. Involving patients, and physicians in the discussion are necessary to understand which symptoms matter to patients. At the same time, it is important to consider if these symptoms can be objectively measured and what is a meaningful change in measurement that reflects treatment benefit.

Once it is clear what endpoints need to be captured, the right device can be selected. The device technology needs to be verified (measurement errors, variances, etc.) and the device also needs to be validated for the specific use (reliability; accuracy and precision compared to gold standard or independent measurements). An observational study is required to ensure the suitability of the device before deploying it in a trial.

#### ***Diseases that can be tracked with digital biomarkers***

Heart disease and diabetes measurements are common application areas for sensor-based devices. However, digital BMs could have the most impact in monitoring CNS diseases since it gives us the opportunity to measure symptoms that were largely intractable until now. Various sensor devices are available for tracking several aspects of health such as activity, heart rate, blood glucose and even sleep, breath, voice, and temperature. Most smartphones are equipped with several sensors that can perform the various motion, sound and light based tests. In addition, the smartphone can be used for psychological tests or to detect finger motions through the touchscreen. These measures can be used in various combinations to predict the health aspects or symptoms required.

Digital BMs can have several applications beyond clinical trials, for example in diagnostics—to identify patients affected by a disease. However, the most interesting application is in digital therapeutics where the device/app can be used to help the treatment like insulin dose adjustment or to monitor/treat substance abuse or addiction. Digital BMs present a big opportunity for measuring endpoints in a remote, objective and unbiased manner that was largely difficult until now. However, there are still several challenges that need to be considered before developing and deploying them to measure endpoints in clinical trials.

#### **The conundrum of biomarker strategy**

There is a wrong notion that by the time a BM is discovered and validated; it is too late to affect the decision-making process. The real question is whether the chosen BM is: (1) intrinsically related to the pathogenesis of a disease; and (2) whether it is reliable and adequate for decision-making. It has been reported that building computer models can transform potential BM into clinically meaningful tests. However, on several occasions when scientists [47] attempted to import data from the literature, they found that the diagnostic criteria used to assess BMs accuracy were vague or based on un-validated BMs.

Identifying BMs that can be translated from animal models to humans is also challenging [48]. While inhibiting an enzyme in an animal model may be effective, this may not be the case in humans. This is either because the pathway has diverged or humans have some compensatory mechanisms. A treatment may change a BM, but this may be irrelevant to a specific disease. Therefore, a true BM must be intrinsically linked to the pathogenesis of the disease. A drug should treat a disease, not a BM.

Without understanding the pathogenesis of a disease, it is difficult to figure out what is the right BM to be used in clinical studies. Once a BM is identified, it is difficult to understand whether it is associated with a specific disease or multiple diseases or if it is a reflection of poor health. For instance, if you are studying potential BMs for Systemic Lupus Erythematosus (SLE) or Alzheimer's Disease (AD), the same set of BMs keeps emerging as potential differentiators. A growing body of evidence indicates that SLE is associated with increased risk of cognitive impairment and dementia [49]. The real question is, however, whether those specific BMs would be able to differentiate SLE from AD. Otherwise, the plethora of BMs that has been generated would be irrelevant.

Pharmaceutical companies are obsessed with the idea that a BM needs to be validated before it can be used for decision-making. Unfortunately, there are no clear-cut criteria to date identifying which BM should be validated. The rigor on how to use a BM to kill a compound relies entirely on the discretion of pharmaceutical companies. The risk of using the wrong BM or selecting the wrong set of BMs may lead to the wrong decision of dumping a good drug because the adopted BM strategy was evaluated inaccurately. To overcome this problem, pharmaceutical companies tend to rely their decision-making process on a long list of BMs (very often too many). This is based on the notion that clusters of variables can be used to differentiate responders from non-responders. The risk of utilizing a long list of BMs is not only costly but also to make the data difficult to be interpreted. The best solution to this problem is to select a strategy

that selects a few BMs with complementary predictive properties.

In the last few years, the FDA has pressured pharmaceuticals to shift the paradigm towards PM, thus targeting diagnostics and treatments based on patient-stratification. This has prompted everyone in the pharmaceutical field to translate molecular profiles into effective treatments, thus impacting: (i) prevention; (ii) early detection; (iii) use of animal or in silico models to facilitate the prediction of success by increasing efficacy and minimizing toxicity and (iv) computational biology to create new synergies between discovery and drug development.

#### **Computational biology and bioinformatics to aid biomarker development**

There is a need to develop novel computer-aided algorithms and methodologies for pattern recognition, visualization, and classification of distribution metrics for interpreting large sets of data coming from high-throughput molecular profiling studies. This is where the bioinformatics and computational biology play a critical role in linking biological knowledge with clinical practice: they are the interface between the clinical development process of drug target and BM discovery and development.

Computational biology uses computational tools and machine learning for data mining, whereas bioinformatics applies computing and mathematics to the analysis of biological datasets to support the solution of biological problems. Bioinformatics plays a key role in analyzing data generated from different 'omics' platforms annotating and classifying genes/pathways for target identification and disease association.

The goal of bioinformaticians is to use computational methods to predict factors (genes and their products) using: (1) a combination of mathematical modeling and search techniques; (2) mathematical modeling to match and analyze high-level functions; and (3) computational search and alignment techniques to compare new biomolecules (DNA, RNA, protein, metabolite, etc.) within each functional 'omics' platform. Combination of this and patient datasets are then used to generate hypotheses.

Bioinformatics and computational biology enable fine tuning of hypotheses [50]. These fields often require specialized tools and skills for data exploration, clustering, regression and supervised classification [51, 52], pattern recognition and selection [53], and development of statistical filtering or modeling strategies and classifiers including neural networks or support vector machines.

The integration of clinical and 'omics' data sets has allowed the exploitation of available biological data such as functional annotations and pathway data [54–56].

Consequently, this has led to the generation of prediction models of disease occurrence or responses to therapeutic intervention [51, 57].

However, the use of high throughput "omics" technologies and large sample sizes have generated massive amounts of data sets and their analyses have become a major bottleneck requiring sophisticated computational and statistical methods and skill sets to analyze them [9].

#### **The role of modeling and simulation to support information-based medicine**

Modeling and simulation (M&S) can accelerate drug development and reduce costs significantly [58]. It relies on a feedback loop leading to the production of more relevant compounds to feed into the development cycle. M&S begins with a new data set, such as BMs to link bench to bedside, thus generating a feedback loop with the drug development cycle. Once the right data is available, investigators can test hypotheses to understand the molecular factors contributing to disease and devising better therapies and simulating different study designs before testing the drug candidate in a clinical trial.

The utility of this approach was shown by Roche AG receiving approval for a combination drug (PEGASYS) for the treatment of hepatitis C. The approach used a variety of factors, including the genotype of the virus and the weight of the patient to select the proper dose for a subset of patients. Pfizer was also pioneering this approach for Neurontin (gabapentin). This drug was approved for a variety of neuropathic pain disorders, including post-herpetic neuralgia. Despite those examples, many companies have not yet fully embraced this approach and are still struggling with modeling and simulation tools, due to poor integration of separate data sets. The tools developed for data integration do not communicate well with each other since they rely on data that are in separate databases. Based on this, it will be difficult to include M&S as an integral part of the development process, unless companies integrate their systems more seamlessly. All the odds speak in favor of the fact that industries are adopting standard data formats and managing structured (data in databases) and unstructured data (documents) sets. As a result, the outcome of translating drug development into clinical practice will be more efficient.

Using pharmacogenomic data, M&S can help us to unravel critical safety issues. FDA has started to recognize with the Critical Path initiative the value of M&S as an important part of the CRADA in 2006 (US Food and Drug Administration, "Challenge and Opportunity on the Critical Path to New Medical Products").

The goal of CRADA is to develop software to support CDISC data formats that can link to other FDA databases

and which can ultimately conduct modeling and simulation. This data will ultimately be applied to the end of Phase IIa revision to make a go or no-go decision.

#### **Machine learning and artificial intelligence can improve precision medicine**

The recent big data revolution, accompanied with the generation of continuously collected large data set from various molecular profiling (genetic, genomic, proteomic, epigenomic and others) efforts of patient samples by the development and deployment of wearable medical devices (e.g. wearable watches) and mobile health applications, and clinical outcome data has enabled the biomedical community to apply artificial intelligence (AI) and machine learning algorithms to vast amounts of data. These technological advancements have created new research opportunities in predictive diagnostics, precision medicine, virtual diagnosis, patient monitoring, and drug discovery and delivery for targeted therapies. These advancements have awoken the interests of academic, industry researchers, and regulatory agencies alike and are already providing new tools to physicians.

An example is the application of precision immunoprofiling by image analysis and artificial intelligence to biology and disease. This was demonstrated in a recent paper where the authors used immunoprofiling data to assess immuno-oncology biomarkers, such as PD-L1 and immune cell infiltrates as predictors of patient's response to cancer treatment [5]. Through spatial analysis of tumor-immune cell interactions, multiplexing technologies, machine learning, and AI tools these authors demonstrated the utility of pattern-recognition in large and complex datasets and deep learning approaches for survival analysis [5].

Essentially, we are using genetics, epigenetics, genomics, proteomics, and other molecular profiling data to inform biology, which we then are evaluating progressively backward using clinical, cellular, and in vitro assays for the discovery of novel targets, pathways, and BMs. Using this plethora of data and data on drugs, we are in a position to come up with candidate drugs faster that most likely work as compared to rational drug design. The goal for human exploratory data would be to aggregate data across the entire medical ecosystem, and give it to third parties to analyze. The pharmaceutical industry could then use AI to build models or to surface patterns—connecting with the patient outcome data—to provide insights into potential benefits to patients. To accomplish this, it is going to take academia, government, and industry—society at large to make better use of human exploratory data. Up to date, the only way to streamline access to human exploratory data is if patients consent, so part of the solution is patient empowerment.

A recent publication [59] highlights the potential utility of AI in cancer diagnostics. Scientists re-trained an off-the-shelf Google deep learning algorithm to identify the most common types of lung cancers with 97% accuracy that even identified altered genes driving abnormal cell growth. To accomplish this, scientists fed Inception v3 slide images supplied by The Cancer Genome Atlas, a database consisting of images of cancer histopathology data and the associated diagnostic annotations. This type of AI has been used to identify faces, animals, and objects in pictures uploaded to servers portal (i.e. Google's online services) has proven useful at diagnosing the disease before, including diabetic blindness and heart conditions. The researchers found the AI performed almost as well as experienced pathologists when it was used to distinguish between adenocarcinoma, squamous cell carcinoma, and normal lung tissue. Intriguingly, the program was trained to predict the 10 most commonly mutated genes in adenocarcinoma and found that six of them—STK11, EGFR, FAT1, SETBP1, KRAS, and TP53—can be predicted from pathology images, with AUCs from 0.733 to 0.856 as measured on a held-out population. The genetic changes identified by this study often cause the abnormal growth seen in cancer and they can change a cell's shape and interactions with its surroundings, providing visual clues for automated analysis.

In another study, researchers used machine learning and retrospectively identified multiple factors that underlie cancer immunotherapy success which potentially allows better target immunotherapy treatment to those who will benefit [60]. To generate their computer model, researchers analyzed data (measured mutations and gene expression in the tumor and T cell receptor (TCR) sequences in the tumor and peripheral blood in urothelial cancers treated with anti-PD-L1) from 21 patients with bladder cancer from a clinical trial dataset of urothelial cancers from Snyder et al. [61] with a uniquely rich data set that captured information about tumor cells, immune cells, and patient clinical and outcome data. Instead of modeling the clinical response of each patient directly, researchers modeled the response of each patient's immune system to anti PDL-1 therapy and used the predicted immune responses to stratify patients based on expected clinical benefit. Their computer model identified key features associated with a specific response to the drug (i.e. PD-L1 inhibitor) and applied 36 different features-multi-modal data set into their machine learning algorithm and allowed the algorithm to identify patterns that could predict increases in potential tumor-fighting immune cells in a patient's blood after treatment. The machine learning algorithm identified 20 features. When they analyzed these features as a panel, they were able to describe 79 percent of the variation in patient immune

responses. This suggested that the comprehensive set of features collected and analyzed for these patients may predict the patient immune response with high accuracy. However, if the researchers excluded any one of the three categories from the model (tumor data, immune cell data or patient clinical data) the algorithm can no longer predict immune response with high accuracy and confidence (the model could only predict at most 23 percent of the variation). Authors concluded that integrative models of immune response may improve our ability to predict the patient response to immunotherapy. However, this study only analyzed a small set of patient data (it only incorporated data from 21 patients, which is far too few to be predictive for the general population) and requires validation of this approach in a larger cohort of patients.

We also recently used a similar machine learning approach that enabled us to identify multiple factors that underlie short-term intensive insulin therapy (IIT) therapy success early in the course of type 2 diabetes which potentially allowed better targeted treatment to those patients who will benefit the most [23]. For that, we developed a model that could accurately predict the response to short-term intensive insulin therapy which provided insight into molecular mechanisms driving such response in humans. We selected a machine learning approach based on the random forests (RF) method, which implements an out-of-bag (“bagging”) technique to monitor error and ensure unbiased prediction with reduced risk of overfitting. For our analysis, the RF algorithm was implemented using the “randomForestpackage” in the R environment. As reported by [62], “by using bagging in tandem with random feature selection, the out-of-bag error estimate is as accurate as using a test set of the same size as the training set. Therefore, using the out-of-bag error estimate removes the need for a set aside test set.” In conclusion, our study identified potential responders to IIT (a current limitation in the field) and provided insight into the mechanism of pathophysiologic determinants of the reversibility of pancreatic islet beta-cell dysfunction in patients with early type 2 diabetes.

The advancements in digital health opportunities have also arisen numerous questions and concerns for the future of biomedical research and medical practice especially when it comes to reliability of AI-driven diagnostic tools, the impact of these tools on clinical practice and patients; vulnerability of algorithms to bias and unfairness, and ways to detect and improve the bias and unfairness in machine learning algorithms [63].

In summary, we hope that the AI program in a not too distant future helps to identify or even predict mutations instantly, avoiding the delays imposed by genetic tests, which can take weeks to confirm the presence of mutations. These findings suggest that AI and machine

learning models can assist pathologists in the detection of cancer subtype or gene mutations in an efficient and expeditious way.

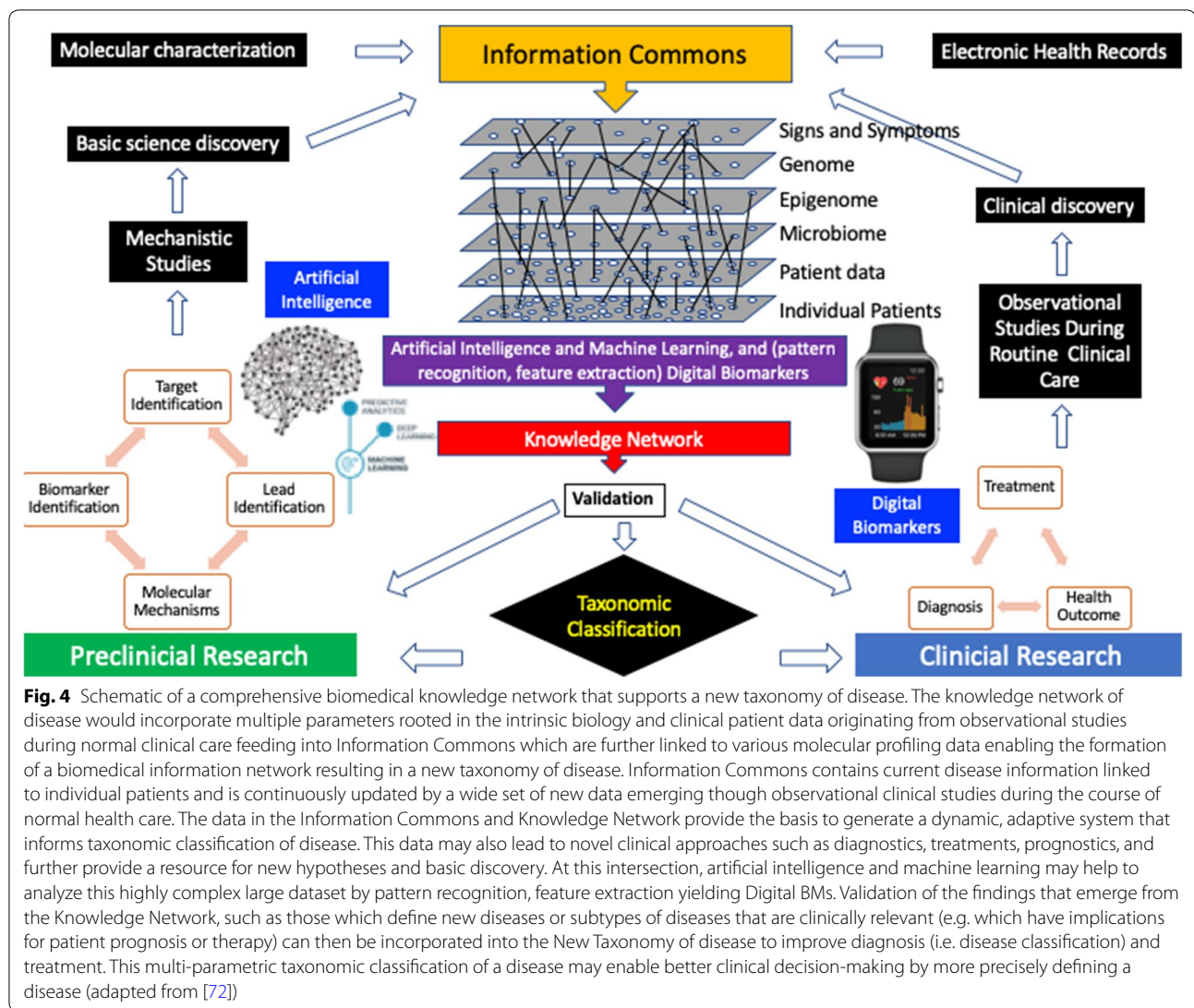
#### **Deep phenotyping—linking physiological abnormalities and molecular states—from bedside to bench**

The analysis of phenotype plays a key role in medical research and clinical practice towards better diagnosis, patient stratification, and selection of best treatment strategies. In biology “phenotype” is defined as the physical appearance or biochemical characteristic of an organism as a result of the interaction between its genotype and the environment “Deep phenotyping” is defined as the precise and comprehensive analysis of phenotypic abnormalities in which the individual components of the phenotype (taking a medical history or a physical examination, diagnostic imaging, blood tests, psychological test, etc., in order to make the correct diagnosis) have been observed and described [64]. However, to understand the pathogenesis of a disease, several key points must be considered, such as the spectrum of complications, classification of patients into more homogeneous subpopulations that differ with respect to disease susceptibility, genetic and phenotypic subclasses of a disease, family history of disease, duration of disease, or to the likelihood of positive or adverse response to a specific therapy.

The concept of “PM” which aims to provide the best available medical care for each individual, refers to the stratification of patients into more homogeneous subpopulations with a common biological and molecular basis of disease, such that strategies developed from this approach is most likely to benefit the patients [Committee on the Framework for Developing a New Taxonomy of Disease, 2011]. A medical phenotype comprises not only the abnormalities described above but also the response of a patient to a specific type of treatment. Therefore, a better understanding of the underlying molecular factors contributing to disease and associated phenotypic abnormalities requires that phenotype is linked to molecular profiling data.

Therefore, deep phenotyping, combined with advanced molecular phenotypic profiling such as genetics and genomics including Genome-wide association studies (GWAS), epigenetics, transcriptomics, proteomics, and metabolomics, with all their limitations, enables the construction of causal network models (Fig. 4) in which a genomic region is proposed to influence the levels of transcripts, proteins, and metabolites. This takes advantage of the relative (*i.e.* the function of regulatory RNAs and epigenetic modifications on phenotype) unidirectional flow of genetic information from DNA variation to phenotype.





As discussed by Schadt et al. [65] the relationships between various physiological phenotypes (e.g. physiological traits) and molecular phenotypes (e.g. DNA variations, variations in RNA transcription levels, RNA transcript variants, protein abundance, or metabolite levels) together constitute the functional unit which must be examined to understand the link to disease and strata of more homogeneous population representing the phenotype. All this can accelerate the identification of disease subtypes with prognostic or therapeutic implications, and help to develop better treatment strategies. Therefore, phenotypic analysis bears great importance for elucidating the physiology and pathophysiology of networks at the molecular and cellular level because it provides clues about groups of genes, RNAs, or proteins that constitute pathways or modules, in which dysfunction can lead to phenotypic consequences. Several recent studies

have shown the utility of correlating phenotypes to features of genetic or cellular networks on a genome scale [66–69]. The emerging field of “Knowledge Engineering for Health” proposes to link the research to the clinic by using deep phenotypic data to enable research based on the practice and outcomes of clinical medicine which consequently lead to decision making in stratified and PM contexts [70].

#### The knowledge network of disease

As illustrated in Fig. 4, and further discussed in the literature [71] a knowledge network of disease should integrate multiple datasets and parameters to yield a taxonomy heavily embedded in the intrinsic biology of disease. Despite physical signs and symptoms are the overt manifestations of disease, symptoms are often non-specific and rarely identify a disease with confidence and they are

not as objective and not quantitative. In addition, a number of diseases— such as different types of cancer, cardiovascular disease, and HIV infection are asymptomatic in early stages. As a result, diagnosis based on traditional “signs and symptoms” alone carries the risk of missing opportunities for prevention, or early intervention.

On the other hand, advances in liquid biopsies, which analyze cells, DNA, RNA, proteins, or vesicles isolated from the blood as well as microbiomes have gained particular interest for their uses in acquiring information reflecting the biology of health and disease state. Biology-based BMs of disease such as genetic mutations, protein, metabolite BMs, epigenetic alterations of DNA, alterations in gene expression profiles, circulating miRNAs, cell-free DNAs, exosomes, and other biomolecules have the potential to be precise descriptors of disease.

When multiple BMs are used in combination with conventional clinical, histological, and laboratory findings, they often are a more accurate, sensitive, specific for the precise description and classification of disease.

In the near future, it is anticipated that comprehensive molecular profiling and characterization of healthy persons and patients will take place routinely as a normal part of health care even as a preventive measure prior to the appearance of disease, thus enabling the collection of data on both healthy and diseased individuals on a grander scale. The ability to conduct molecular characterizations on both non-affected and disease affected tissues would enable monitoring of the development and natural history of many diseases.

## Summary

The drug development is a challenging long process with many obstacles on the way. Though several strategies have been proposed to tackle this issue, there is a general consensus that a better use of BMs, omics data, AI and machine learning will accelerate the implementation of a new medical practice that will depart from the widely spread concept “one drug fits all”.

In conclusion, drug developers must combine traditional clinical data with patients’ biological profile including various omics-based datasets to generate an “information-based” model that utilizes complex datasets to gain insight into disease and facilitate the development of more precise, safer, and better-targeted therapies for a more homogeneous patient population.

## Review criteria

Publicly available information such as PubMed and Internet were used for the literature review. We focused on identifying articles published on the use of multiple technologies for the discovery and development of clinically relevant BMs, omics platforms, and other

relevant topics in the subject area. The research was restricted to the most recent studies in this field and all research was limited to human studies published in English.

## Abbreviations

AD: Alzheimer’s Disease; Ag: antigen; AI: artificial intelligence; BMs: biomarkers; BETi: inhibitors of bromodomain and extra terminal proteins; Carbo: carboplatin; Cf-DNA: cell-free DNA; cf-RNA: cell-free RNA CSF1: colony stimulating factor 1; CDX: companion diagnostics; CFM: cyclophosphamide; CTC: circulating tumor cells; CTLA-4: cytotoxic T-lymphocyte-associated antigen 4; CTTI: The Clinical Trials Transformation Initiative; DNA: deoxyribonucleic acid; E.g.: *exempli gratia*; Etc.: etcetera; FDA: The Food and Drug Administration; FLS: fibroblast-like synoviocytes; GWAS: genome-wide association study; HDAC: histone deacetylase; HMA: hypomethylating agents; ICGC: International Cancer Genome Consortium (ICGC); IDO: indoleamine 2,3-dioxygenase; I.E.: *Id Est*; IIT: intensive insulin therapy; LAG-3: lymphocyte-activation gene 3; LN: lymph nodes; MDSC: myeloid-derived suppressor cells; MHC: major histocompatibility complex; M&S: modeling and simulation; miRNAs: microRNAs; MS: multiple sclerosis; OA: osteoarthritis; IO: immuno-oncology; P13K: phosphoinositide 3-kinase; PD-1: programmed cell death-1; PD-L1: programmed death-ligand 1; PM: precision medicine; RA: rheumatoid arthritis; RF: random forests; RNA: ribonucleic acid; SLE: systemic lupus erythematosus; STING: stimulator of interferon genes; TCR: T cell receptor (TCR); TIM3: T cell immunoglobulin and mucin domain 3; TLR: toll-like receptor; TILs: tumor-infiltrating lymphocytes; TME: tumor microenvironment; Treg: regulatory T cells; Wnt: wingless/integrated 1.

## Authors’ contributions

Both authors participated in the conceptualization of the thematic of this review article and contributed equally to create and critically review the manuscript. In addition, AS created the Figure and the Table that was adopted from publicly available dataset. Both authors read and approved the final manuscript.

## Author details

<sup>1</sup> Department of Pathology and Laboratory Medicine, Division of Biology and Medicine, Brown University, Providence, RI 02903, USA. <sup>2</sup> Fox Chase Cancer Center, Temple University Temple Health, Philadelphia, PA 19111, USA. <sup>3</sup> School of Cancer and Pharmaceutical Sciences, Faculty of Life Sciences & Medicine, King’s College London, London SE1 8WA, UK.

## Acknowledgements

We thank our affiliated institutes making this publication possible. We apologize to the many authors and colleagues whose works are not cited due to limited space.

## Competing interests

The authors declare that they have no competing interests.

## Availability of data and materials

Not applicable.

## Consent for publication

Not applicable.

## Ethics approval and consent to participate

This article does not contain any studies with human participants or animals performed by any of the authors. Informed consent was not required for the preparation of this review article as it used secondary sources only.

## Funding

No funding involved in the preparation of this article.

## Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 11 February 2019 Accepted: 28 March 2019  
Published online: 05 April 2019

## References

- Seyhan A, Carini C. Biomarkers for drug development: the time is now. Carini C, Menon S, Chang M, editors. Clinical and statistical considerations in personalized medicine. Chapman & Hall: CRC Press; 2014. p. 16–41.
- Seyhan AA. Biomarkers in drug discovery and development. *Eur Biopharm Rev*. 2010;1:19–25.
- Cesano A, Warren S. Bringing the next Generation of immuno-oncology biomarkers to the clinic. *Biomedicines*. 2018;6:14.
- Blank CU, Haanen JB, Ribas A, Schumacher TN. The “cancer immunogram”. *Science*. 2016;352:658–60.
- Koelzer VH, Sirinukunwattana K, Rittscher J, Mertz KD. Precision immunoprofiling by image analysis and artificial intelligence. *Virchows Arch*. 2018. <https://doi.org/10.1007/s00428-018-2485-z>.
- Lee HJ, Seo AN, Kim EJ, Jang MH, Kim YJ, Kim JH, Kim SW, Ryu HS, Park IA, Im SA, et al. Prognostic and predictive values of EGFR overexpression and EGFR copy number alteration in HER2-positive breast cancer. *Br J Cancer*. 2014;112:103.
- Carini C, Seyhan A. From isolation to integration: a systems biology approach for the discovery of therapeutic targets and biomarkers. Barker KB, Menon S, Agostino R, Xu S, Jin B, eds. *Biosimilar Clinical development: scientific considerations and new methodologies*. 2016. p. 2.
- Selleck MJ, Senthil M, Wall NR. Making meaningful clinical use of biomarkers. *Biomark Insights*. 2017;12:1177271917715236.
- Poste G. Bring on the biomarkers. *Nature*. 2011;469:156–7.
- Burke HB. Predicting clinical outcomes using molecular biomarkers. *Biomark cancer*. 2016;8:89–99.
- Butterfield LH, Disis ML, Fox BA, Lee PP, Khleif SN, Thurin M, Trinchieri G, Wang E, Wigginton J, Chaussabel D, et al. A systematic approach to biomarker discovery; preamble to “the iSBTC-FDA taskforce on immunotherapy biomarkers”. *J Transl Med*. 2008;6:81.
- Lang L. High clinical trials attrition rate is boosting drug development costs. *Gastroenterology*. 2004;127:1026.
- Moutinho-Ribeiro P, Macedo G, Melo SA. Pancreatic cancer diagnosis and management: has the time come to prick the bubble? *Front Endocrinol*. 2019;9:799.
- de la Rica L, Urquiza JM, Gomez-Cabrero D, Islam AB, Lopez-Bigas N, Tegner J, Toes RE, Ballestar E. Identification of novel markers in rheumatoid arthritis through integrated analysis of DNA methylation and microRNA expression. *J Autoimmun*. 2013;41:6–16.
- Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B, Liew A, Khalili H, Chandrasekaran A, Davies LR, et al. TRAF1-C5 as a risk locus for rheumatoid arthritis—a genome-wide study. *N Engl J Med*. 2007;357:1199–209.
- Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, de Bakker PI, Le JM, Lee HS, Batliwalla F, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. *N Engl J Med*. 2007;357:977–86.
- Alevizos I, Illei GG. MicroRNAs in Sjogren’s syndrome as a prototypic autoimmune disease. *Autoimmun Rev*. 2010;9:618–21.
- Hezova R, Slaby O, Faltejskova P, Mikulkova Z, Buresova I, Raja KR, Hodek J, Ovesna J, Michalek J. microRNA-342, microRNA-191 and microRNA-510 are differentially expressed in T regulatory cells of type 1 diabetic patients. *Cell Immunol*. 2010;260:70–4.
- Seyhan AA, Nunez Lopez YO, Xie H, Yi F, Mathews C, Pasarica M, Pratley RE. Pancreas-enriched miRNAs are altered in the circulation of subjects with diabetes: a pilot cross-sectional study. *Sci Rep*. 2016;6:31479.
- Seyhan AA. microRNAs with different functions and roles in disease development and as potential biomarkers of diabetes: progress and challenges. *Mol BioSyst*. 2015;11:1217–34.
- Nunez Lopez YO, Garufi G, Seyhan AA. Altered levels of circulating cytokines and microRNAs in lean and obese individuals with prediabetes and type 2 diabetes. *Mol BioSyst*. 2016;13:106–21.
- Nunez Lopez YO, Pittas AG, Pratley RE, Seyhan AA. Circulating levels of miR-7, miR-152 and miR-192 respond to vitamin D supplementation in adults with prediabetes and correlate with improvements in glycemic control. *J Nutr Biochem*. 2017;49:117–22.
- Lopez YON, Retnakaran R, Zinman B, Pratley RE, Seyhan AA. Predicting and understanding the response to short-term intensive insulin therapy in people with early type 2 diabetes. *Mol Metab*. 2019;20:63–78.
- Nunez Lopez YO, Coen PM, Goodpaster BH, Seyhan AA. Gastric bypass surgery with exercise alters plasma microRNAs that predict improvements in cardiometabolic risk. *Int J Obes (Lond)*. 2017;41:1121–30.
- Nunez Lopez YO, Garufi G, Pasarica M, Seyhan AA. Elevated and correlated expressions of miR-24, miR-30d, miR-146a, and SFRP-4 in human abdominal adipose tissue play a role in adiposity and insulin resistance. *Int J Endocrinol*. 2018;2018:7351902.
- Pachori AS, Madan M, Nunez Lopez YO, Yi F, Meyer C, Seyhan AA. Reduced skeletal muscle secreted frizzled-related protein 3 is associated with inflammation and insulin resistance. *Obesity (Silver Spring)*. 2017;25:697–703.
- Du C, Liu C, Kang J, Zhao G, Ye Z, Huang S, Li Z, Wu Z, Pei G. MicroRNA miR-326 regulates TH-17 differentiation and is associated with the pathogenesis of multiple sclerosis. *Nat Immunol*. 2009;10:1252–9.
- Lindberg RL, Hoffmann F, Mehling M, Kuhle J, Kappos L. Altered expression of miR-17-5p in CD4+ lymphocytes of relapsing-remitting multiple sclerosis patients. *Eur J Immunol*. 2010;40:888–98.
- Nakasa T, Miyaki S, Okubo A, Hashimoto M, Nishida K, Ochi M, Asahara H. Expression of microRNA-146 in rheumatoid arthritis synovial tissue. *Arthritis Rheum*. 2008;58:1284–92.
- Pauley KM, Satoh M, Chan AL, Bubb MR, Reeves WH, Chan EK. Upregulated miR-146a expression in peripheral blood mononuclear cells from rheumatoid arthritis patients. *Arthritis Res Ther*. 2008;10:R101.
- Padgett KA, Lan RY, Leung PC, Lleo A, Dawson K, Pfeiff J, Mao TK, Coppel RL, Ansari AA, Gershwin ME. Primary biliary cirrhosis is associated with altered hepatic microRNA expression. *J Autoimmun*. 2009;32:246–53.
- Tang Y, Luo X, Cui H, Ni X, Yuan M, Guo Y, Huang X, Zhou H, de Vries N, Tak PP, et al. MicroRNA-146A contributes to abnormal activation of the type I interferon pathway in human lupus by targeting the key signaling proteins. *Arthritis Rheum*. 2009;60:1065–75.
- Dai Y, Sui W, Lan H, Yan Q, Huang H, Huang Y. Comprehensive analysis of microRNA expression patterns in renal biopsies of lupus nephritis patients. *Rheumatol Int*. 2009;29:749–54.
- Sonkoly E, Wei T, Janson PC, Saaf A, Lundeberg L, Tengvall-Linder M, Norstedt G, Alenius H, Homey B, Scheynius A, et al. MicroRNAs: novel regulators involved in the pathogenesis of psoriasis? *PLoS ONE*. 2007;2:e610.
- Chatzikiriakidou A, Voulgari PV, Georgiou I, Drosos AA. miRNAs and related polymorphisms in rheumatoid arthritis susceptibility. *Autoimmun Rev*. 2012;11:636–41.
- Alevizos I, Illei GG. MicroRNAs as biomarkers in rheumatic diseases. *Nat Rev Rheumatol*. 2010;6:391–8.
- Alsaleh G, Suffert G, Semaan N, Juncker T, Frenzel L, Gottenberg JE, Sibilia J, Pfeffer S, Wachsmann D. Bruton’s tyrosine kinase is involved in miR-346-related regulation of IL-18 release by lipopolysaccharide-activated rheumatoid fibroblast-like synoviocytes. *J Immunol*. 2009;182:5088–97.
- Nakamachi Y, Kawano S, Takenokuchi M, Nishimura K, Sakai Y, Chin T, Saura R, Kurosaka M, Kumagai S. MicroRNA-124a is a key regulator of proliferation and monocyte chemoattractant protein 1 secretion in fibroblast-like synoviocytes from patients with rheumatoid arthritis. *Arthritis Rheum*. 2009;60:1294–304.
- Seyhan AA, Varadarajan U, Choe S, Liu W, Ryan TE. A genome-wide RNAi screen identifies novel targets of neratinib resistance leading to identification of potential drug resistant genetic markers. *Mol BioSyst*. 2012;8:1553–70.
- Seyhan AA, Varadarajan U, Choe S, Liu Y, McGraw J, Woods M, Murray S, Eckert A, Liu W, Ryan TE. A genome-wide RNAi screen identifies novel targets of neratinib sensitivity leading to neratinib and paclitaxel combination drug treatments. *Mol BioSyst*. 2011;7:1974–89.
- Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. *Nature*. 2017;541:321–30.
- Nagarsheth N, Wicha MS, Zou WP. Chemokines in the cancer micro-environment and their relevance in cancer immunotherapy. *Nat Rev Immunol*. 2017;17:559–72.
- Torous J, Andersson G, Bertagnoli A, Christensen H, Cuijpers P, Firth J, Haim A, Hsin H, Hollis C, Lewis S, et al. Towards a consensus around standards for smartphone apps and digital mental health. *World Psychiatry*. 2019;18:97–8.

44. Boulos MN, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. *Online J Public Health Inform.* 2014;5:229.
45. Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, Koelsche C, Sahm F, Chavez L, Reuss DE, et al. DNA methylation-based classification of central nervous system tumours. *Nature.* 2018;555:469–74.
46. Moran S, Martinez-Cardus A, Sayols S, Musulen E, Balana C, Estival-Gonzalez A, Moutinho C, Heyn H, Diaz-Lagares A, de Moura MC, et al. Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. *Lancet Oncol.* 2016;17:1386–95.
47. Diniz BS, Pinto Junior JA, Forlenza OV. Do CSF total tau, phosphorylated tau, and beta-amyloid 42 help to predict progression of mild cognitive impairment to Alzheimer's disease? A systematic review and meta-analysis of the literature. *World J Biol Psychiatry.* 2008;9:172–82.
48. Cavagnaro JA. Preclinical safety evaluation of biotechnology-derived pharmaceuticals. *Nat Rev Drug Dis.* 2002;1:469.
49. Zhao Z, Rocha NP, Salem H, Diniz BS, Teixeira AL. The association between systemic lupus erythematosus and dementia. A meta-analysis. *Dement Neuropsychol.* 2018;12:143–51.
50. Azuaje F, Devaux Y, Wagner D. Computational biology for cardiovascular biomarker discovery. *Brief Bioinform.* 2009;10:367–77.
51. Camargo A, Azuaje F. Identification of dilated cardiomyopathy signature genes through gene expression and network data integration. *Genomics.* 2008;92:404–13.
52. Frank E, Hall M, Trigg L, Holmes G, Witten IH. Data mining in bioinformatics using Weka. *Bioinformatics.* 2004;20:2479–81.
53. Saeyns Y, Inza I, Larranaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics.* 2007;23:2507–17.
54. Deschamps AM, Spinale FG. Pathways of matrix metalloproteinase induction in heart failure: bioactive molecules and transcriptional regulation. *Cardiovasc Res.* 2006;69:666–76.
55. Camargo A, Azuaje F. Linking gene expression and functional network data in human heart failure. *PLoS ONE.* 2007;2:e1347.
56. Ginsburg GS, Seo D, Frazier C. Microarrays coming of age in cardiovascular medicine: standards, predictions, and biology. *J Am Coll Cardiol.* 2006;48:1618–20.
57. Ideker T, Sharan R. Protein networks in disease. *Genome Res.* 2008;18:644–52.
58. McGee P. Modeling success with in silico tools. *Drug Discov Dev.* 2005;8:23–8.
59. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, Moreira AL, Razavian N, Tsirigos A. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med.* 2018;24:1559.
60. Leiserson MDM, Syrgkanis V, Gilson A, Dudik M, Gillett S, Chayes J, Borgs C, Bajorin DF, Rosenberg JE, Funt S, et al. A multifactorial model of T cell expansion and durable clinical benefit in response to a PD-L1 inhibitor. *PLoS ONE.* 2018;13:e0208422.
61. Snyder A, Nathanson T, Funt SA, Ahuja A, Buros Novik J, Hellmann MD, Chang E, Aksoy BA, Al-Ahmadie H, Yusko E, et al. Contribution of systemic and somatic factors to clinical response and resistance to PD-L1 blockade in urothelial cancer: an exploratory multi-omic analysis. *PLoS Med.* 2017;14:e1002309.
62. Breiman L. Random forests. *Mach Learn.* 2001;45:5–32.
63. National Academies of Sciences E. Medicine: artificial intelligence and machine learning to accelerate translational research: proceedings of a workshop—in brief. Washington, DC: The National Academies Press; 2018.
64. Robinson PN. Deep phenotyping for precision medicine. *Hum Mutat.* 2012;33:777–80.
65. Schadt EE. Molecular networks as sensors and drivers of common human diseases. *Nature.* 2009;461:218–23.
66. Feldman I, Rzhetsky A, Vitkup D. Network properties of genes harboring inherited disease mutations. *Proc Natl Acad Sci USA.* 2008;105:4323–8.
67. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabasi AL. The human disease network. *Proc Natl Acad Sci USA.* 2007;104:8685–90.
68. Robinson PN, Kohler S, Bauer S, Seelow D, Horn D, Mundlos S. The human phenotype ontology: a tool for annotating and analyzing human hereditary disease. *Am J Hum Genet.* 2008;83:610–5.
69. Kohler S, Bauer S, Horn D, Robinson PN. Walking the interactome for prioritization of candidate disease genes. *Am J Hum Genet.* 2008;82:949–58.
70. Beck T, Gollapudi S, Brunak S, Graf N, Lemke HU, Dash D, Buchan I, Diaz C, Sanz F, Brookes AJ. Knowledge engineering for health: a new discipline required to bridge the "ICT gap" between research and healthcare. *Hum Mutat.* 2012;33:797–802.
71. Council NR. Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease. Washington, DC: The National Academies Press; 2011.
72. In toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease. Washington (DC); 2011: The National Academies Collection: Reports funded by National Institutes of Health.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

