Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Pharmacology

How scientific literature analysis yields innovative therapeutic hypothesis through integrative iterations

Francois Iris, Athanasios Beopoulos and Manuel Gea

It is becoming generally accepted that the current diagnostic system often guarantees, rather than diminishes, disease heterogeneity. In effects, syndrome-dominated conceptual thinking has become a barrier to understanding the biological causes of complex, multifactorial diseases characterized by clinical and therapeutic heterogeneity. Furthermore, not only is the flood of currently available medical and biological information highly heterogeneous, it is also often conflicting. Together with the entire absence of functional models of pathogenesis and pathological evolution of complex diseases, this leads to a situation where illness activity cannot be coherently approached and where therapeutic developments become highly problematic. Acquisition of the necessary knowledge can be obtained, in parts, using *in silico* models produced through analytical approaches and processes collectively known as 'Systems Biology'. However, without analytical approaches that specifically incorporate the facts that all that is called 'information' is not necessarily useful nor utilisable and that all information should be considered as a priori suspect, modelling attempts will fail because of the much too numerous conflicting and, although correct in molecular terms, physiologically invalid reports. In the present essay, we suggest means whereby this body of problems could be functionally attacked and describe new analytical approaches that have demonstrated their efficacy in alleviating these difficulties.

Address
Bio-Modeling Systems, Tour CIT, 3 Rue de l'Arrivée, 75015, Paris, France

Corresponding author: Iris, Francois (francois.iris@bmsystems.org)

Current Opinion in Pharmacology 2018, **42**:62–70

This review comes from a themed issue on **New technologies**

Edited by **Antoine Bril**, **Oliver Nosjean**, and **Patrick Genissel**

https://doi.org/10.1016/j.coph.2018.07.005

Background

Drug development is primarily a problem of data integration and knowledge management. Knowing the potential targets of a molecule and the functions of these targets is one thing. Understanding the physiological mechanisms that must be targeted and the manner in which they must be manipulated to have a therapeutic impact is quite another. Thus, success in therapeutic development largely depends upon the coherent manipulation of a physiological system in its pathological context and not upon the manipulation of a target in a molecular setting. Yet, identification of the presence of a given pathology is largely based upon the symptoms presented by any given patient. These symptoms, together with the results of medical and biological tests, are then utilised to reach a medical diagnostic. In practice, most experienced physicians utilise the pattern recognition method to identify the clinical problem. Theoretically, a given pattern of tests results and symptoms within a given local population context can be directly associated with a given therapy, even without a definite decision regarding what is the actual disease [1].

Hence, the vast majority of complex disorders are defined by a number of symptoms that can differ considerably between affected individuals with respect to their presence, frequency, severity and topology. Indeed, within a population context, different individuals may present similar symptoms for totally different physiological reasons just as they can present different symptoms for very similar physiological reasons. However, the compromise that constitutes the pattern recognition method, which primarily relies upon the information available to the physician, carries a substantial risk of misdiagnosis, confusing different pathologies which actually require different therapies. This is most evident in the context of complex pathologies [2–5].

Furthermore, heterogeneity in symptoms complicates the search for the aetiology of complex diseases and the mechanisms for their treatment. In effects, the current diagnostic system often guarantees, rather than diminishes, disease heterogeneity and current syndrome-dominated conceptual thinking has become a barrier to understanding the biological causes of a wide variety of diseases characterized by clinical and therapeutic heterogeneity such as muscular dystrophies [6], mitochondrial dysfunctions [7,8], retinal degenerative diseases [9,10], thyroid pathologies [11], autoimmune [12] and neurological diseases [13,14], metabolic [15••,16] and psychiatric disorders [17], and so on.

This leads to an untenable situation that precludes coherent therapeutic developments since it effectively

prevents defining what could constitute valid biological, clinical and therapeutic biomarkers.

The issues of biomarkers in drug development

Biomarkers are at the roots of evidence-based medicine (who should be treated, how and with what) and without valid biomarkers, not only advances in better targeted therapies will remain limited but treatments will also remain largely empirical. Furthermore, biomarkers for improved prediction and monitoring of disease and toxicology mechanisms are needed to control the high clinical failure rates among new compounds [18,19].

But, in the absence of clear pathophysiological understanding, the maturity and utility of safety-related biomarkers varies very significantly among target organ systems [20••,21].

A 'biomarker' is typically defined as a laboratory measurement that reflects the activity of a disease process or the responses to a therapeutic intervention. But the goals of therapeutic interventions are twofold: (1) better symptomatic therapies, and (2) treatments that slow disease progression or delay disease onset. This necessarily leads to a second class of biomarkers, known as 'clinical endpoints', that are not measured for the purpose of detecting clinical benefit but for their reflection of the underlying pathological process [22]. In essentially all cases, these markers must quantitatively correlate, either directly or inversely, with disease progression. Taking into account the state of our current understanding of pathological processes, this literally opens a Pandora box. In the context of functionally heterogeneous disorders, there might be as many biomarkers as there are affected individuals. Hence, the much sought-after 'gold standard biomarkers' for a set of individuals affected by a common disease remains an unattainable goal [23].

In attempts to circumvent these issues, a third class of biomarkers has been put forward, the so-called 'surrogate markers'. This object is defined as a laboratory measurement or physical sign that is used in therapeutic trials as a substitute for a clinically meaningful endpoint and is a direct measure of how a patient feels, functions, or survives and is thus expected to predict the effects of a therapy [24]. Hence, the major difference between a biomarker and a surrogate marker is that a biomarker is a 'candidate' surrogate marker, whereas a surrogate marker is a test used, and taken, as a measure of the effects of a specific treatment.

However, drugs development must necessarily proceed through pre-clinical studies carried out on laboratory animal models, usually inbred rodent strains, which present the apparent symptomatology of the human pathology being addressed but rarely its actual physiological basis. Not only these animal models often amount to mere caricatures of the human pathology, but the results of the drugs development assays are also interpreted according to their effects upon the animal model's symptomatology, hence an all too frequent inadequacy with respect to the human physiopathology with ensuing clinical trial failures or drug withdrawals from the market.

As a result, strong efforts are now being devoted to the search for combinatorial biomarkers, generated through high content screening, and in particular high content in situ proteomics and imaging technologies, to be used in the industry to screen for toxic side effects of drug candidates and to identify appropriate patient populations [25] in the hope that this will support the knowledge-based decision-making process by providing crucial information on functional biology [26,27]. In doing so, it is assumed that a given symptomatically defined disorder (semiology) necessarily implicates restricted sets of physiological mechanisms, some of which must eventually be shared among affected patients, irrespective of their environments. It thus becomes a matter of screening a sufficiently large number of patients to hence identify relevant markers or combinations thereof.

These approaches therefore identify 'biomarkers' as a function of their statistical occurrence and not in terms of their physiological relevance. The phenomena that give rise to disease and responses to treatments heterogeneities, the very reason behind the search in the first place, are entirely ignored. Furthermore, individuals affected by a severe disease often present a variety of concurrently induced/associated disorders (comorbidities), some of which may remain under-diagnosed and their prevalence under-rated [28•,29,30]. If it is accepted that a pathology must necessarily leave traces of its presence under the form of biomarkers as defined above, then the concurrent presence of another pathology, whether clinically recognised or not, must also necessarily do so.

Thus, far from helping to resolve the issues generated by the syndrome-dominated vision, this further worsens an already difficult situation, particularly in the case of heterogeneous disorders. These shortcomings have for net results to reiterate previous costly mistakes, albeit under a different form. Not only statistical effects are expected to compensate for lack of knowledge, but an additional flaw is now being introduced. Differences in physical environments are implicitly considered as having little impact upon the biological mechanisms associated with defined semiologies [31•,32]. This directly leads to a highly deleterious situation already experienced by the industry in the past.

Indeed, in order to increase drug development successes, it was found necessary to significantly increase the size and the scope of clinical trials. The main reasons for this were associated with the phenomenon of functional