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

disease heterogeneity coupled with the impossibility to define trustworthy exclusion criteria applicable at recruitment levels [33]. The results are a massive escalation in costs with very limited returns in terms of successes [34•].

The biological reality

Semiology is a poor indicator of similarities or differences in associated physiological processes. Not only functionally different disease states can present similar semiologies, but a disease with well identified biological causes often present considerable variability in both semiology and outcome among affected patients, including between siblings [35,36]. Here, both physical and sociological environments play important roles [37–39]. All the more so since they also affect epigenetic mechanisms and subsequent disease susceptibility without change in primary DNA sequence [40,41,42••]. All these effects are particularly apparent in the numerous cases of discordances among monozygotic twins for a wide variety of disorders, many of which are regarded as having significant genetic backgrounds [43,44].

The outcome is that physiopathologically different forms of a disease much too frequently fail to be recognised as such while heterogeneous presentations of a same physiopathology lead to differential diagnosis [45,46]. This, inescapably, leads to sample misclassifications that can reach very significant proportions [47]. Hence, given the biological reality briefly described above, it seems hardly reasonable to expect that useful potential clinical and /or prognostic and/or therapeutic biomarkers could be identified on the basis of their statistical occurrence. To have the least chance of success, knowledge of what to search for, where, when and why appears to be a necessity.

Biological functions results from interactions between integrative and non-linear mechanisms

Phenotypes and behaviour depend on the integrated effects of multiple signalling pathways and molecules, genetic polymorphisms, epigenetic marking dynamics and epistatic phenomena as well as environmental stimuli affecting post-translational processing. This is true for all biological systems, from individual cells all the way to organisms.

Interventions upon such systems must take into account the fact that biological functions result from dynamic, integrative and non-linear processes subject to discontinuities. Indeed, it has become obvious that the mere state assessment of components and metabolites in a living system does not reveal in a predictable and reliable manner the activity of pathways and circuits that they comprise [48•,49]. Furthermore, many organs and biological functions are eminently affected by time (the act of being actively alive), in terms of both structural anatomy as well as networks interactions [50,51]. Functional evolution over time (ageing) implies numerous switches in gene expression patterns [52,53]. Interpreted in terms of functional physiology, this leads to the inescapable conclusion that, in any individual, the functional situation implemented when time-point 'B' will be reached may be radically different from what it was at time-point 'A'. The numerous reported cases of spontaneous remission in a variety of sever pathological conditions [54–57] may well be striking examples of this.

Hence, it is not merely a matter of what proteins (components) are expressed and to what level, but rather of what other potential interaction partners are present and in what state. It is qualitative aspects that are important here, and not quantitative considerations. A protein deemed 'physiologically important' may be entirely absent or constitutively non-functional without producing deleterious phenotypic effects. This is amply demonstrated by the numerous such instances observed in knockout mice [58]. Furthermore, the deleterious effects of inactivating mutations affecting a given protein can often be compensated by inactivating mutations simultaneously affecting another protein [59] or by corrective mechanisms its functional absence induce [60••,61,62].

The needs for systemic biological models and the role of systems biology

The necessarily brief survey above provides some indication of the density, extent and enormous complexity of the integrative effects, resulting from dynamic interactions between a multitude of biological factors and interconnected pathways, leading to most medical conditions requiring therapeutic interventions [63••,64]. This makes simplistic pharmacological approaches ineffective when not actually damaging [65,66]. Indeed, in most instances, pharmacological interventions take the form of monotherapies, each targeting a very limited, if not a single set of physiological mechanisms. Besides the numerous undesirable effects these approaches trigger, which are often at the root of pre-clinical and, more damagingly, clinical failures [67•], *de novo* and acquired resistances to treatment are widespread [68–70]. It is becoming evident that, for the vast majority of complex disorders, therapeutic approaches that simultaneously target multiple pathways are urgently required [71–73].

The challenge for future drug development will thus be to devise pharmacological approaches that reflect the overall pathophysiological and biological processes that must be affected to produce defined therapeutic responses whilst retaining the simplicity and robustness required for routine pre-clinical and clinical testing.

Appropriate pathophysiological understanding will therefore be necessarily required for both therapeutic targeting purposes as well as for the identification of biomarkers that are physiologically relevant for clinical (patients

stratification), prognostic (disease status) and therapeutic (response/efficacy/toxicity) purposes [74••].

The analytical and biological complexities to be mastered are such that none of this can be achieved using reductionist (classical) approaches. To have the least chance of success, we must develop predictive functional models sufficiently detailed so as to enable the precise identification, in mechanistic terms, of events leading to pathological consequences, hence identifying the mechanisms and the markers associated with these events together with the modes of intervention most likely to prevent or alleviate the problems. Acquisition of the necessary knowledge can be obtained, in parts, using *in silico* theoretical models produced through analytical approaches and processes collectively known as 'Systems Biology'. Systems biology addresses the need to shift from a component-based reductionist view of biology to a system-wide perspective. It can be described as a global analysis of how all components in a biological system interact to determine its phenotype. Although the definitions may vary, systems biology can usually be characterized as interdisciplinary, iterative, computationally intensive, and information greedy.

Systems biology applied to the discovery and validation of novel therapeutic approaches

The usual approaches to systems biology are characterized by

(1) their dependence upon relevant quantitative data arising from multiple targeted experimental interrogations in an iterative interplay between experimentation and modelling, with
(2) the aim of elucidating how the molecular components of a living system determine its phenotype by exploring their dynamic interplays as well as their interactions with the environment.

Thus, this approach interprets biological phenomena as dynamic processes, the mechanisms and consequences of which depend on the behaviour of components that constitute the living entity studied.

While this requirement may be reasonably fulfilled in the context of well-defined mechanisms (e.g. the Rab5–Rab7 toggle and cut-out switches in the conversion of early endosomes into late endosomes) [75], it can hardly be contemplated when addressing complex, heterogeneous and often obscure pathophysiological mechanisms. Indeed, when dealing with issues addressing complex human pathologies, data on relevant molecular and intercellular dynamics is seldom available to an extent and a range likely to sustain classical modelling approaches [76••]. Here, rather than adamantly adhering to the established principles attached to the most frequently used approaches to systems biology, attempts could be made to reconsider the problem under a different light.

One such alternative approach to systems analysis could be based on the exploitation of the tremendous amount of information contained in the existing published scientific literature and in highly heterogeneous publicly accessible databases rather than on homogeneous datasets arising from specifically targeted investigations.

However, the matter is much more arduous than might be anticipated.

'Information' is a double-edged tool to be manipulated with caution

It is indeed certain that, if predictive models of complex and heterogeneous diseases are to ever be constructed, enormous masses of information originating from a multitude of biological investigations and encompassing an enormous functional complexity will have to be integrated.

Although daunting in amplitude, if approached coherently, the flood of highly heterogeneous, and often conflicting, information that currently hampers most biological fields can become an invaluable tool. But this tool must be approached and manipulated with extreme caution.

Most of the huge amounts of currently available information arose from reductionist approaches that, in attempts to compensate for the enormous, and often insurmountable, experimental difficulties presented by *in vivo* systems, utilised *in vitro* experimentations on material frequently far removed from functional physiological reality. Hence, the enormous diversity of information obtained in association with most physiological networks represents a highly distorted, compounded view of the various modulations that can potentially affect each such network, albeit without distinction of actual *in vivo* physiological relevance. Furthermore, due to the wide diversity of biological systems that gave rise to this information, and the often complete lack of physiological compatibilities between experimental systems, the information thereby generated is necessarily always (1) incomplete, to an unknown extent; (2) biased, in unknown manners and to an unknown extent; and (3) erroneous, to an unknown extent, and this irrespective of the domain addressed. Indeed, it is currently considered that an estimated 85% of research resources are wasted, many published research findings being false or exaggerated [77–79,80••].

As a consequence, 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 fails because the much too numerous conflicting and physiologically invalid reports will inevitably lead to an accumulation of analytically crippling inconsistencies.

However, the above very real difficulties certainly do not mean that the scientific literature cannot be used for the implementation of systems-based analyses. These difficulties simply highlight the fact that to coherently utilise the highly heterogeneous information available, novel analytical approaches become necessary.

Changing the analytical paradigm

By focusing mainly on chemical and physical processes with the expectation that living systems can be fully explained from this engineer's perspective, the classical approach to systems biology assumes bottom-up causation, from molecular dynamics to cellular/tissue behaviour.

However, the stability of a living system lies in its homeostatic capacity to re-establish itself.

In a living system, the outcome does not crucially depend on strictly predefined operations of the parts. Rather, the structure of the whole determines the operation of the parts. Indeed, almost all homeostatic processes are complex context-dependent entities to which genes make a necessary, but only partial, contribution.

In such systems, homeostasis proceeds on the basis of functional loops wherein on-going events tell local physiological contexts how to evolve, contexts tell components how to behave and components tell future events how to arise, and so on.

In other words, specific biological events do not occur because they are fated to. They occur because other events could not arise.

It follows that analyses in terms of biological components and functions now become irrelevant. What become necessary are event-driven analytical approaches.

In addition to this, the intrinsic value of any 'information' is only relative. It can be profoundly modified by other, indirectly linked information as well as by the contexts it can be attached to.

Thus, both the available information and the biological processes to be considered are characterized by heavily context-dependent attributes.

Therefore, whatever event-driven analytical approach is implemented, it must also be 'relativistic'. That is, all available information must be treated on the basis of a negative selection procedure. What can be identified as false can in turn be used to discover what could be true, provided that a heuristic and event-driven analytical procedure is implemented.

Heuristics can be characterized as a problem solving approach evaluating each step in a process, searching for satisfactory solutions rather than for optimal solutions, using all available qualitative information. Thus, heuristic modelling starts from accumulated knowledge to produce, via iterative integrations, a model capable of describing the biological events and the mechanisms that generated the observed biological phenomenon and predict the modifications they will sustain in association with any given intervention.

The above considerations constitute the functional basis on which the CADI (computer-assisted deductive integration) analytical procedure was developed.

The logic behind this model-building approach (Figure 1) does not assume functional linearity and the components of a model do not incorporate solely what is known. Indeed, since this approach relies upon strict and systematic implementation of negative selection of hypotheses, models arising from this procedure contain elements that had hitherto never been described but cannot be refuted by current knowledge and/or available biological data, thereby generating novel understanding.

Here, heuristic and mathematical modelling, far from being antagonistic, are complementary. Heuristic modelling plays the role of an architect (defines the nature, the structure, the functionalities and the contextual constraints of the system under study) whereas mathematical modelling, to be implemented at a later stage, plays the role of an engineer (reveals the dynamics and robustness of the structures within the system while defining the set of parameters sufficient to give rise to similar or very different phenotypes).

Although the models arising from this analytical approach cannot, by any means, be regarded as biologically true in the absolute, they do represent a 'least biased' and detailed view of the mechanisms potentially associated with a given physiological state and/or governed by the biological components under consideration, with precise indications of the means whereby these could be manipulated together with the identification of the most relevant biomarkers and their significance.

In other words, these models clearly indicate what should be biologically observed in a given context, where, when, how and why. Data newly gathered from investigations aiming to test the model's validity can then be re-injected in the model-building procedure, allowing rapid and efficient correction of the model, thereby rapidly leading to a clear and factual understanding of the biological

Figure 1

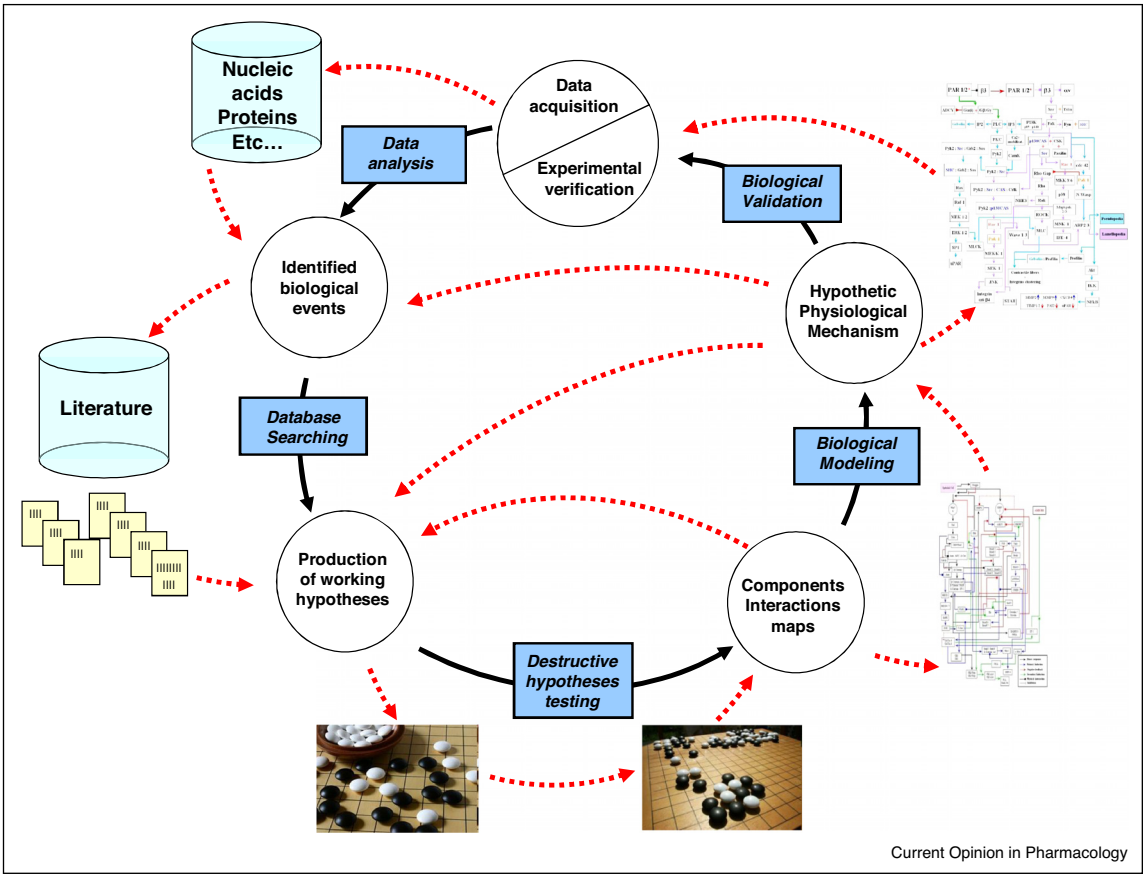

Diagram depicting of the model-building procedure. The circular shape of the overall process, embodied by labelled black arrows, indicates the iterative nature of the entire analytical procedure where each main analytical step is represented by a labelled disc. Since a circle does not have a defined beginning, the model-building procedure may be initiated from any given step in the process. The dotted red arrows indicate the internal iterative loops within the overall process. The go board on the lower left (black & white tokens on a checker board) represents the distribution of untested working hypotheses directly generated from the literature while its lower right counterpart represents the distribution of hypotheses that have resisted all destruction attempts. These undestroyed hypotheses are then merged to produce interaction maps that are in turn merged to produce a hypothetical physiological model. During each of these two steps, numerous novel working hypotheses suggesting hitherto undescribed biological events are being generated. These are, in turn, subjected to the iterative negative (destructive) selection procedure. This is represented by dotted red arrows linking 'interaction maps' and 'hypothetical physiological model' (lower and upper right discs) with 'identified events' and 'production of working hypotheses' (upper and lower left discs). Hence, this model-building process involves multiple levels of internal crosscheck procedures designed to eliminate any hypothesis that is not directly as well as indirectly supported by multiple data intersects (scientific literature and publicly accessible databases). The model resulting from this process is then subjected to direct experimental evaluation and the new data thus generated can be in turn integrated into the model, correcting errors and mis-directions.

processes under investigation while now providing a therapeutically directly exploitable representation of the biological reality addressed.

Conclusion

The models produced through this approach belong to the class of non-mathematical, holistic and heuristic models. By abandoning the usual symptomatic point of view, they perform the necessary switch to diseases analyses from a causal standpoint. They constitute detailed maps of inter-cellular and/or intra-cellular mechanisms associated with a biological state and they describe the progression mechanisms of a pathological process as well as the functional characteristics of pathological state versus control.

This alternative model-building approach, which synergistically associates algorithmics and heuristics, results in a situation where biological investigations proceed at a rapid pace (months instead of years) with a hitherto unachievable success rate.

Indeed, this approach has repeatedly proven its efficacy in the discovery of (i) hitherto unsuspected biological mechanisms, pathways and interactions directly associated with phenotypic transitions *in vivo* (be they pathological or

developmental) [81,82,83•], (ii) patent protected novel therapeutic approaches in fields ranging from oncology to neurodegenerative and infectious diseases [84–88], and (iii) the development of novel, patent protected technologies [89].

Furthermore, when applied to neurodegenerative disorders, this approach was selected by the EU's DG Research as one of three examples of 'state-of-the-art' in systems biology that benefit to medicine [90] and was granted an industrial "Best Practice Award" by The Cambridge Health Tech Institute (USA) [91].

However, it is important to realize that such models can only be an approximation of biological reality. Furthermore, the more complex the reality attached to a model, the coarser the model will be. It is therefore indispensable that such a model be confronted to the biological reality it claims to represent.

Conflict of interest statement

Nothing declared.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as

- • of special interest
- •• of outstanding interest

1. Hay MC, Weisner TS, Subramanian S *et al.*: **Harnessing experience: exploring the gap between evidence-based medicine and clinical practice**. *J Eval Clin Pract* 2008, **14**:707-713.
2. Busby LP, Courtier JL, Glastonbury CM: **Bias in radiology: the how and why of misses and misinterpretations**. *Radiographics* 2018, **38**:236-247.
3. Rakha EA, Ahmed MA, Aleskandarany *et al.*: **Diagnostic concordance of breast pathologists: lessons from the National Health Service Breast Screening Programme Pathology External Quality Assurance Scheme**. *Histopathology* 2017, **70**:632-642.
4. Weng QY, Raff AB, Cohen JM *et al.*: **Costs and consequences associated with misdiagnosed lower extremity cellulitis**. *JAMA Dermatol* 2016 http://dx.doi.org/10.1001/jamadermatol.2016.3816.
5. Solomon AJ, Bourdette DN, Cross AH *et al.*: **The contemporary spectrum of multiple sclerosis misdiagnosis: a multicenter study**. *Neurology* 2016, **87**:1393-1399.
6. Klein A, Clement E, Mercuri E, Muntoni F: **Differential diagnosis of congenital muscular dystrophies**. *Euro J Paediat Neurol* 2008, **12**:371-377.
7. Siciliano G, Pasquali L, Mancuso M, Murri L: **Molecular diagnostics and mitochondrial dysfunction: a future perspective**. *Expert Rev Mol Diagn* 2008, **8**:531-549.
8. Wong LJ: **Diagnostic challenges of mitochondrial DNA disorders**. *Mitochondrion* 2007, **7**:45-52.
9. Lenarduzzi S, Vozzi D, Morgan A *et al.*: **Usher syndrome: an effective sequencing approach to establish a genetic and clinical diagnosis**. *Hear Res* 2015, **320**:18-23.
10. den Hollander AI, Roepman R, Koenekoop RK, Cremers FP: **Leber congenital amaurosis: genes, proteins and disease mechanisms**. *Prog Retin Eye Res* 2008, **27**:391-419.
11. Scognamiglio T: **C cell and follicular epithelial cell precursor lesions of the thyroid**. *Arch Pathol Lab Med* 2017, **141**:1646-1652.
12. Harada K, Hiep NC, Ohira H: **Challenges and difficulties in pathological diagnosis of autoimmune hepatitis**. *Hepatol Res* 2017, **47**:963-971.
13. Post B, Speelman JD, de Haan RJ, group CA-S: **Clinical heterogeneity in newly diagnosed Parkinson's disease**. *J Neurol* 2008, **255**:716-722.
14. Sagna A, Gallo JJ, Pontone GM: **Systematic review of factors associated with depression and anxiety disorders among older adults with Parkinson's disease**. *Parkinsonism Relat Disord* 2014, **20**:708-715.
15. •• Fenger M, Linneberg A, Werge T, Jorgensen T: **Analysis of heterogeneity and epistasis in physiological mixed populations by combined structural equation modelling and latent class analysis**. *BMC Genet* 2008, **9**:43.
Gives a very good account of the reasons why syndrome-dominated thinking impedes understanding diseases characterized by clinical and therapeutic heterogeneity.
16. Bermudez VJ, Rojas E, Toledo A *et al.*: **Single-nucleotide polymorphisms in adiponectin, AdipoR1, and AdipoR2 genes: insulin resistance and type 2 diabetes mellitus candidate genes**. *Am J Ther* 2013, **20**:414-421.
17. Keshavan MS, Nasrallah HA, Tandon R: **Schizophrenia, 'Just the Facts" 6. Moving ahead with the schizophrenia concept: from the elephant to the mouse**. *Schizophr Res* 2011, **127**:3-13.
18. Frantz S: **Pharma faces major challenges after a year of failures and heated battles**. *Nat Rev Drug Discov* 2007, **6**:5-7.
19. Kotecki N, Penel N, Awada A: **How to emerge from the conservatism in clinical research methodology?** *Curr Opin Oncol* 2017, **29**:400-404.
20. •• Rolan P, Danhof M, Stanski D, Peck C: **Current issues relating to drug safety especially with regard to the use of biomarkers: a meeting report and progress update**. *Eur J Pharm Sci* 2007, **30**:107-112.
Important for understanding the needs for effective biomarkers and the difficulties attached.
21. Twerenbold R, Reichlin T, Reiter M, Muller C: **High-sensitive cardiac troponin: friend or foe?** *Swiss Med Wkly* 2011, **141**:w13202.
22. Gainor JF, Longo DL, Chabner BA: **Pharmacodynamic biomarkers: falling short of the mark?** *Clin Cancer Res* 2014, **20**:2587-2594.
23. Nallagangula KS, Nagaraj SK, Venkataswamy L, Chandrappa M: **Liver fibrosis: a compilation on the biomarkers status and their significance during disease progression**. *Future Sci* 2017, **4**:FSO250.
24. Parast L, McDermott MM, Tian L: **Robust estimation of the proportion of treatment effect explained by surrogate marker information**. *Stat Med* 2016, **35**:1637-1653.
25. Koop R: **Combinatorial biomarkers: from early toxicology assays to patient population profiling**. *Drug Discov Today* 2005, **10**:781-788.
26. Christians U, Schmitz V, Schoning W *et al.*: **Toxicodynamic therapeutic drug monitoring of immunosuppressants: promises, reality, and challenges**. *Ther Drug Monit* 2008, **30**:151-158.
27. Slovin SF: **Pitfalls or promise in prostate cancer immunotherapy — which is winning?** *Cancer J* 2008, **14**:26-34.
28. • Lam CS: **Diabetic cardiomyopathy: an expression of stage B heart failure with preserved ejection fraction**. *Diab Vasc Dis Res* 2015, **12**:234-238.
Gives a good example of asymptomatic and under-diagnosed comorbidity with significant physiological consequences.
29. Gonzalez-Duarte A, Garcia-Ramos GS, Valdes-Ferrer SI, Cantu-Brito C: **Clinical description of intracranial hemorrhage associated with bleeding disorders**. *J Stroke Cerebrovasc Dis* 2008, **17**:204-207.

30. Apte M, McGwin G, Jr, Vila LM *et al.*: **Associated factors and impact of myocarditis in patients with SLE from LUMINA, a multiethnic US cohort (LV)**. *Rheumatology (Oxford)* 2008, **47**:362-367.

31. • Jog NR, James JA: **Biomarkers in connective tissue diseases**. *J Allergy Clin Immunol* 2017, **140**:1473-1483.
A good account of the lack of correlation between statistical occurrence and physiological relevance for biomarkers in multiples disorders.

32. Gonzalez-Angulo AM, Morales-Vasquez F, Hortobagyi GN: **Overview of resistance to systemic therapy in patients with breast cancer**. *Adv Exp Med Biol* 2007, **608**:1-22.

33. Patel SK, Velkoska E, Burrell LM: **Emerging markers in cardiovascular disease: where does angiotensin-converting enzyme 2 fit in?** *Clin Exp Pharmacol Physiol* 2013, **40**:551-559.

34. • Kola I: **The state of innovation in drug development**. *Clin Pharmacol Ther* 2008, **83**:227-230.
Good analysis of the reasons behind the current inefficiency of drugs development.

35. Moreno R, Lardennois C, Drouin-Garraud V *et al.*: **Prenatal revelation of Niemann–Pick disease type C in siblings**. *Acta Paediatr* 2008, **97**:1136-1139.

36. Coan AC, Cendes F: **Understanding the spectrum of temporal lobe epilepsy: contributions for the development of individualized therapies**. *Exp Rev Neurother* 2013, **13**:1383-1394.

37. Navarro MC, Sosa M, Saavedra P *et al.*: **Anthropometric and gynaecological history according to the socioeconomic status of postmenopausal women: poverty and the menopause**. *Menopause Int* 2010, **16**:12-17.

38. Law C: **Early growth and chronic disease: a public health overview**. *Matern Child Nutr* 2005, **1**:169-176.

39. Zilhao NR, Smit DJ, Boomsma DI, Cath DC: **Cross-disorder genetic analysis of tic disorders, obsessive–compulsive, and hoarding symptoms**. *Front Psychiatry* 2016, **7**:120.

40. Dolinoy DC, Jirtle RL: **Environmental epigenomics in human health and disease**. *Environ Mol Mutagen* 2008, **49**:4-8.

41. McGowan PO, Szyf M: **Environmental epigenomics: understanding the effects of parental care on the epigenome**. *Essays Biochem* 2010, **48**:275-287.

42. •• Swathy B, Banerjee M: **Understanding epigenetics of schizophrenia in the backdrop of its antipsychotic drug therapy**. *Epigenomics* 2017, **9**:721-736.
Good analysis of the influence of epigenetic mechanisms upon disease susceptibility, clinical progression and effects of treatment.

43. Haque FN, Gottesman II, Wong AH: **Not really identical: epigenetic differences in monozygotic twins and implications for twin studies in psychiatry**. *Am J Med Genet C Semin Med Genet* 2009, **151C**:136-141.

44. Kan KJ, Dolan CV, Nivard MG *et al.*: **Genetic and environmental stability in attention problems across the lifespan: evidence from the Netherlands twin register**. *J Am Acad Child Adolesc Psychiatry* 2013, **52**:12-25.

45. Marchioni E, Tavazzi E, Bastianello S: **Non-multiple sclerosis recurrent demyelinating disorders: an ongoing debate**. *Brain* 2010, **133**:e150.

46. Yohe S, Yeh IT: **''Missed" diagnoses of phyllodes tumor on breast biopsy: pathologic clues to its recognition**. *Int J Surg Pathol* 2008, **16**:137-142.

47. Jellinger KA: **Morphological substrates of Parkinsonism with and without dementia: a retrospective clinico-pathological study**. *J Neural Transm Suppl* 2007, **72**:91-104.

48. • Holmes WE, Angel TE, Li KW, Hellerstein MK: **Dynamic proteomics: in vivo proteome-wide measurement of protein kinetics using metabolic labeling**. *Methods Enzymol* 2015, **561**:219-276.
Important to understand the functional role of dynamic processes, including protein kinetics, in disease pathogenesis and clinical progression.

49. Sauer U: **Metabolic networks in motion: 13C-based flux analysis**. *Mol Syst Biol* 2006, **2**:62.

50. Hirjak D, Meyer-Lindenberg A, Kubera KM, Thomann PA, Wolf RC: **Motor dysfunction as research domain in the period preceding manifest schizophrenia: a systematic review**. *Neurosci Biobehav Rev* 2018, **87**:87-105.

51. Proix T, Spiegler A, Schirner M, Rothmeier S, Ritter P, Jirsa VK: **How do parcellation size and short-range connectivity affect dynamics in large-scale brain network models?** *Neuroimage* 2016, **142**:135-149.

52. Moskalev AA, Aliper AM, Smit-McBride Z, Buzdin A, Zhavoronkov A: **Genetics and epigenetics of aging and longevity**. *Cell Cycle* 2014, **13**:1063-1077.

53. Maric NP, Svrakic DM: **Why schizophrenia genetics needs epigenetics: a review**. *Psychiatr Danub* 2012, **24**:2-18.

54. Vriends N, Bolt OC, Kunz SM: **Social anxiety disorder, a lifelong disorder? A review of the spontaneous remission and its predictors**. *Acta Psychiatr Scand* 2014, **130**:109-122.

55. Trumpf J, Becker ES, Vriends N, Meyer AH, Margraf J: **Rates and predictors of remission in young women with specific phobia: a prospective community study**. *J Anxiety Disord* 2009, **23**:958-964.

56. Shima Y, Nakanishi K, Hama T *et al.*: **Spontaneous remission in children with IgA nephropathy**. *Pediatr Nephrol* 2013, **28**:71-76.

57. Poddubnyy D, Gensler LS: **Spontaneous, drug-induced, and drug-free remission in peripheral and axial spondyloarthritis**. *Best Pract Res Clin Rheumatol* 2014, **28**:807-818.

58. Barbaric I, Miller G, Dear TN: **Appearances can be deceiving: phenotypes of knockout mice**. *Brief Funct Genomic Proteomic* 2007, **6**:91-103.

59. Kobsar I, Oetke C, Kroner A, Wessig C, Crocker P, Martini R: **Attenuated demyelination in the absence of the macrophage-restricted adhesion molecule sialoadhesin (Siglec-1) in mice heterozygously deficient in P0**. *Mol Cell Neurosci* 2006, **31**:685-691.

60. •• Parlato R, Cruz H, Otto C *et al.*: **Effects of the cell type-specific ablation of the cAMP-responsive transcription factor in noradrenergic neurons on locus coeruleus firing and withdrawal behavior after chronic exposure to morphine**. *J Neurochem* 2010, **115**:563-573.
Good example of the corrective mechanisms induced by the constitutive absence of a protein deemed 'physiologically important'.

61. Parlato R, Rieker C, Turiault M, Tronche F, Schutz G: **Survival of DA neurons is independent of CREM upregulation in absence of CREB**. *Genesis* 2006, **44**:454-464.

62. de Caprona MD, Beisel KW, Nichols DH, Fritzsch B: **Partial behavioral compensation is revealed in balance tasked mutant mice lacking otoconia**. *Brain Res Bull* 2004, **64**:289-301.

63. •• Labandeira-Garcia JL, Rodriguez-Perez AI, Garrido-Gil P *et al.*: **Brain renin–angiotensin system and microglial polarization: implications for aging and neurodegeneration**. *Front Aging Neurosci* 2017, **9**:129.
Good account of the extent and complexity of the integrative effects, resulting from dynamic interactions between a multitude of biological factors and interconnected pathways, leading to medical conditions requiring therapeutic interventions.

64. Kas MJ, Modi ME, Saxe MD, Smith DG: **Advancing the discovery of medications for autism spectrum disorder using new technologies to reveal social brain circuitry in rodents**. *Psychopharmacology (Berl)* 2014, **231**:1147-1165.

65. Volpato M, Phillips RM: **Tailoring targeted therapy to individual patients: lessons to be learnt from the development of mitomycin C**. *Cancer Genomics Proteom* 2007, **4**:175-186.

66. Poljsak B, Suput D, Milisav I: **Achieving the balance between ROS and antioxidants: when to use the synthetic antioxidants**. *Oxid Med Cell Longev* 2013, **2013**:956792.

67. • Kleinschnitz C, Meuth SG, Wiendl H: **The trials and errors in MS therapy**. *Int MS J* 2008, **15**:79-90.

Reviews the underlying reasons for discrepancies between theoretical expectations and physiological outcomes leading to drugs clinical failures.

68. Fabi A, Mottolese M, Segatto O: **Therapeutic targeting of ERBB2 in breast cancer: understanding resistance in the laboratory and combating it in the clinic**. *J Mol Med (Berl)* 2014, **92**:681-695.

69. Bergers G, Hanahan D: **Modes of resistance to anti-angiogenic therapy**. *Nat Rev Cancer* 2008, **8**:592-603.

70. Millan MJ, Goodwin GM, Meyer-Lindenberg A, Ove Ogren S: **Learning from the past and looking to the future: emerging perspectives for improving the treatment of psychiatric disorders**. *Eur Neuropsychopharmacol* 2015, **25**:599-656.

71. Petrelli A, Giordano S: **From single- to multi-target drugs in cancer therapy: when aspecificity becomes an advantage**. *Curr Med Chem* 2008, **15**:422-432.

72. Rexer BN, Ghosh R, Arteaga CL: **Inhibition of PI3K and MEK: it is all about combinations and biomarkers**. *Clin Cancer Res* 2009, **15**:4518-4520.

73. Hamon M, Blier P: **Monoamine neurocircuitry in depression and strategies for new treatments**. *Prog Neuropsychopharmacol Biol Psychiatry* 2013, **45**:54-63.

74. •• Ruiz-Cerda ML, Irurzun-Arana I, Gonzalez-Garcia I *et al.*: **Towards patient stratification and treatment in the autoimmune disease lupus erythematosus using a systems pharmacology approach**. *Eur J Pharm Sci* 2016, **94**:46-58.
Highlights the needs for dentification of targets, biomarkers and patient subpopulations with differential response to drugs and the fact that a sparse data situation imposes modelling approaches other than traditional quantitative models based on differential equations.

75. Hu YB, Dammer EB, Ren RJ, Wang G: **The endosomal–lysosomal system: from acidification and cargo sorting to neurodegeneration**. *Transl Neurodegener* 2015, **4**:18.

76. •• Ji Z, Yan K, Li W, Hu H, Zhu X: **Mathematical and computational modeling in complex biological systems**. *Biomed Res Int.* 2017, **2017**:5958321.
Excellent description of mathematical and computational systemic modeling and the challanges posed by complex, heterogeneous and often obscure pathophysiological mechanisms.

77. Wicherts JM: **The weak spots in contemporary science (and how to fix them)**. *Animals (Basel)* 2017, **7**:piiE90.

78. Blatt MR: **When is science 'ultimately unreliable'?** *Plant Physiol* 2016, **170**:1171-1173.

79. Page MJ, McKenzie JE, Kirkham J *et al.*: **Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions**. *Cochrane Database Syst Rev* 2014:MR000035.

80. •• Ioannidis JP: **How to make more published research true**. *PLoS Med* 2014, **11**:e1001747.
Important factual review of the actual usefulness of current scientific literature.

81. Iris F, Gea M, Lampe PH, Santamaria P: **Production and implementation of predictive biological models**. *Med Sci (Paris)* 2009, **25**:608-616.

82. Iris F: **Psychiatric systems medicine: closer at hand than anticipated but not with the expected portrait**. *Pharmacopsychiatry* 2012, **45**:S12-21.

83. • Iris F., Gea M, Lampe P-H, Querleux B, Querleux Bernard: **Heuristic Modelling Applied to Epidermal Homeostasis**. *Computational Biophysics of the Skin*. Pan Stanford Publishing; 2014:461-524.
Overview of the difficulties to be faced, the means whereby they are overcome together with a concrete example addressing the elucidation of an hitherto obscure differential mechanism.

84. Gadal F, Bozic C, Pillot-Brochet C, Malinge S, Wagner S, Le Cam A, Buffat L, Crepin M, Iris F: **Integrated transcriptome analysis of the cellular mechanisms associated with Ha-ras-dependent malignant transformation of the human breast epithelial MCF7 cell line**. *Nucleic Acids Res* 2003, **31**:5789-5804.

85. Gadal F, Starzec A, Bozic C *et al.*: **Integrative analysis of gene expression patterns predicts specific modulations of defined cell functions by estrogen and tamoxifen in MCF7 breast cancer cells**. *J Mol Endocrinol* 2005, **34**:61-75.

86. Turck CW, Iris F: **Proteome-based pathway modelling of psychiatric disorders**. *Pharmacopsychiatry* 2011, **44**:S54-61.

87. Iris F, Filiou M, Turck CW: **Differential proteomics analyses reveal anxiety-associated molecular and cellular mechanisms in cingulate cortex synapses**. *Am J Psych Neurosci* 2014, **2**:25-42.

88. Nussbaumer M, Asara JM, Teplytska L *et al.*: **Selective mitochondrial targeting exerts anxiolytic effects in vivo**. *Neuropsychopharmacology* 2016, **41**:1751-1758.

89. Pouillot F, Blois H, Iris F: **Genetically engineered virulent phage banks in the detection and control of emergent pathogenic bacteria**. *Biosecur Bioterror* 2010, **8**:155-169.

90. European Commission, DG Research, Directorate of Health: From Systems Biology to Systems Medicine 2010. ftp://ftp.cordis.europa.eu/pub/fp7/health/docs/final-report-systems-medicineworkshop_en.pdf. 5-6

91. Cambridge Healthtech Institute: Bio-IT Best Practice Awards 2009. http://www.bio-itworld.com/BioIT_Article.aspx?id = 93536.