Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Biologically grounded scientific methods: The challenges ahead for combating epidemics

Ariel L. Rivas\textsuperscript{a,}\textsuperscript{*}, Almira L. Hoogesteijn\textsuperscript{b}

\textsuperscript{a} University of New Mexico, United States
\textsuperscript{b} Human Ecology, Centro de Investigación y de Estudios Avanzados (CINVESTAV), Merida, Mexico

**ARTICLE INFO**

**Keywords:**
Scientific methods
Method development
Biologically grounded methods
Conceptualization and operationalization
Data structuring

**ABSTRACT**

The protracted COVID 19 pandemic may indicate failures of scientific methodologies. Hoping to facilitate the evaluation and/or update of methods relevant in Biomedicine, several aspects of scientific processes are here explored.

First, the background is reviewed. In particular, eight topics are analyzed: (i) the history of Higher Education models in reference to the pursuit of science and the type of student cognition pursued, (ii) whether explanatory or actionable knowledge is emphasized depending on the well- or ill-defined nature of problems, (iii) the role of complexity and dynamics, (iv) how differences between Biology and other fields influence methodologies, (v) whether theory, hypotheses or data drive scientific research, (vi) whether Biology is reducible to one or a few factors, (vii) the fact that data, to become actionable knowledge, require structuring, and (viii) the need of inter-/trans-disciplinary knowledge integration.

To illustrate how these topics interact, a second section describes four temporal stages of scientific methods: conceptualization, operationalization, validation and evaluation. They refer to the transition from abstract (non-measurable) concepts (such as ‘health’) to the selection of concrete (measurable) operations (such as ‘quantification of anti-virus specific antibody titers’). Conceptualization is the process that selects concepts worth investigating, which continues as operationalization when data-producing variables viewed to reflect critical features of the concepts are chosen. Because the operations selected are not necessarily valid, informative, and may fail to solve problems, validations and evaluations are critical stages, which require inter/trans-disciplinary knowledge integration.

It is suggested that data structuring can substantially improve scientific methodologies applicable in Biology, provided that other aspects here mentioned are also considered. The creation of independent bodies meant to evaluate biologically oriented scientific methods is recommended.

1. Introduction

The COVID-19 pandemic can be viewed as a global educational experience. Never before, so much was published so fast on a single topic: 147,000 peer-reviewed papers were generated within 15 months. Yet, the lessons emerging from this experience do not seem to be globally learned: the fact that the pandemic has lasted so long suggests that the scientific methodology has been inadequate.

Even one and a half years after the beginning of this unprecedented problem, no worldwide dialogue on how to face it has been observed [1]. With some exceptions, there is no international process that brings together researchers, politicians, and citizens under the same roof [2]. Not surprisingly, failures have occurred.

Unsuccessful efforts include, although are not limited to (i) the hypothesis that some levels of ‘herd immunity’ (a number derived from assumptions on the average number of secondary infections induced by a primary case) can protect—a proposition refuted in places where two or more epidemic waves have occurred even with a very high (76%) level of ‘herd immunity’ [3]; (ii) predictions on the number of cases and deaths expected to occur at a certain time point, which have erred by very large amounts [4]; (iii) the assumption that any vaccine can protect even when new viral mutants emerge [5]; (iv) exaggerated optimism regarding treatments that did not materialize [6], and (v) the belief that a pandemic can go away even when there is no explicit policy on diagnostics [7,8].

\textsuperscript{*} Corresponding author.
E-mail addresses: alrivas@ unm.edu (A.L. Rivas), almirahoo@cinvestav.mx (A.L. Hoogesteijn).

https://doi.org/10.1016/j.ymeth.2021.09.001
Received 14 June 2021; Received in revised form 26 August 2021; Accepted 2 September 2021
Available online 4 September 2021
1046-2023/© 2021 Published by Elsevier Inc.
last assumption relates to the fact that no country knows the number and location of non-symptomatic infections because, given their ‘invisibility’, it would be necessary to test, frequently, 100% of the population with tests designed to diagnose. The previous facts suggest failures involving, partially or totally, the ‘scientific method’ (which, most likely, is a plurality, not a singularity), the individuals or agencies expected to conduct such methods, and/or the procedures that rule these interactions.

This review aims at identifying gaps or opportunities of potential relevance in scientific methodologies applied in Biomedicine. This report is divided into two sections. The first section identifies topics that may be relevant in method development. The second section describes temporal aspects of the scientific process. Together, these sections show a possible template for updating scientific methodologies applicable to Biomedicine.

2. Section 1 – The background

2.1. The history of the scientific method: From higher education models to science policy

While partially emphasized by Aristotle, the origin of the scientific method could probably be tracked back to 1597, when Francis Bacon proposed inductivism [9,10]. Originally, scientific research was perceived as a personal activity. Only in the XIX century science started to be institutionalized and, consequently, its professionalization began [11,12].

State-funded scientific research has been associated with new models of Higher Education. While science was ignored in most of the Middle Ages, two models that promoted science emerged a few years after the French Revolution: (i) in France, in 1808, and (ii) in Germany, in 1810 [12,13].

The French model –created by Napoleon- promoted applied sciences, separated teaching from research and appointed faculty members appointments based on examinations. The model Wilhelm von Humboldt created in Germany merged research with teaching, promoted basic research, protected academic freedom, and validated research through a novel process: peer-reviewed publications [12]. Supported by the massive creation of experimental stations, laboratories, and new scientific journals, the German model was partially imitated in the United States through the ‘Land-Grant’ model created in 1862 [13].

Research results, even after World War II, were viewed as a personal ability, not the consequence of a method. The first academic institutionalization of scientific methodologies took place in 1946, when the London School of Economics created the Department of Philosophy, Logic and Scientific Method, and Karl R. Popper became its first member [14].

Soon after, methodology was emphasized in Psychology [15]. A few years earlier, operationalism emerged in Physics -one pillar of contemporary experimental approaches [16].

While emphasized in Economics, Physics and Psychology, methodological research has not been promoted in all fields. Even today, biomedical institutions lack academic units that promote research on biologically-grounded theory and scientific methods [17].

2.2. From explaining to solving problems

In earlier eras, scientific methodologies pursued explanations. They were generated either by inductions based on observations (as proposed by Aristotle and Bacon) or -as proposed by Popper- hypotheses followed by testing and deductions [18]. In contrast, today, problem-solving and problem prevention tend to be prioritized while -due to the abundance of data- hypothesis testing (and, in general, theory) are less emphasized [19,20].

It has been stated that hypotheses should precede data collection because model-driven analyses require data gathered under defined conditions [21] However, the opposite is also possible: data can be assembled before hypotheses are generated [20].

While, in the past, well-defined problems were frequently encountered and usually had simple solutions, in the XXI century problems can be complex, ill-defined, and dynamic [22]. Ill-defined problems may require new cognitive approaches [23].

Earlier versions of scientific methods were oriented to falsify or corroborate prior knowledge. With the current abundance of data, new approaches also focus on discovery [24]. Given that at least 1/3 of all biomedical research publications released over seven decades were generated in the last 5 years [25], it is questionable whether earlier scientific methods - which emerged when most problems were rather simple and the available knowledge was a fraction of the knowledge available today- still apply.

2.3. Educational models and scientific methods

The purpose of higher education has influenced the philosophy of science which, in turn, has shaped scientific methods. When preservation of the status quo and cultural reproduction were the priority, teaching was separated from research and only applied research was promoted. Because such a model, in education, only pursued the approval of examinations, memorization was the priority (prior knowledge was rewarded, even if invalid or not applicable to solve a specific problem), not knowledge generation and use [12].

When creativity and discovery became the center of economics and national survival, learning how to conduct research has turned into a new priority, even in children education [26]. However, when ill-defined (complex and dynamic) problems predominate, the number of possible solutions may be very large and old techniques may be inadequate. In such situations, visually explicit educational strategies may be helpful [27].

Learning activities that generate research-oriented processes for ill-defined problems (not memorizing solutions for clearly defined, simple problems) is now promoted in many countries [28]. They do not offer pre-established solutions. Instead, they promote self-made construction of processes that, through open-ended evaluations, may shed light on whether the novel knowledge is appropriate or should be reconstructed.

2.4. Theory/hypothesis-driven methods, data-driven scientific methods, or both?

Historically, scientific methods were driven by theories or hypotheses. Now they can also be facilitated by data, technologies, new educational strategies and many other sources of knowledge generation [29–32].

Estimated at 130 exabytes (EB) in 2005, the digital universe was around 40,000 EB in 2020 [33]. While, as societies, we are already in the zettabyte era, medicine and scientific research are suspected to be behind -probably at or before the petabyte era [34,35].

While the abundance of data has been contested as a legitimate scientific methodology [17,36], the combinatorial nature of biologic data is also a source of research and learning opportunities [25]. Instead of reducing all alternatives to just two (while assuming that only one is correct), a not self-limiting paradigm may be more defensible, in which both hypothesis testing and hypothesis generating alternatives coexist [10].

2.5. From reductionist to pattern recognition-oriented, non-reductionist scientific methods

Reductionism attempts to explain complexity with only one or a few factor(s) [37]. In addition, a reductionistic method will attempt to fix (or ignore) the environment. Therefore, reductionism does not consider that two or more factors may interact or that time may modify processes and
outcomes. While reductionism assumes that the whole is equal to the sum of the parts, non-reductionism also considers interactions with the environment and, therefore, assumes that the whole is different from the sum of the parts. Reductionism has prevented or delayed the advancement of numerous areas in Biology [38]. Non-reductionist methods are needed when complex and dynamic processes are observed [39,40].

Pattern recognition may provide an alternative to reductionistic methods. The discovery of hidden patterns may reveal critical properties missed by reductionist methods, such as emergence [40,41]. Because the validity of biological methods necessarily depends on the underlying biology, when more than two biological outcomes exist—e.g., no inflammation, early inflammation, late inflammation—binary approaches will necessarily be erroneous because they only consider two alternatives. Hence, methods that emphasize pattern recognition are preferable when the number of possible outcomes is unknown [42–44].

2.6. Beyond falsifiability: Crucial differences between biomedicine and other fields

To develop non-reductionist methods that promote discovery, critical differences between Physics and Biology should be considered. While, in Physics, order is followed by lack of order, the opposite occurs in Biology: living creatures keep constant their ‘order.’ Thus, life seems to be an exception to the second law of thermodynamics [45,46].

Biological data also tend to show ambiguity, i.e., similar values of the same variable may be found at different biological conditions or processes, such as the early and late phases of inflammation. One consequence of biological data ambiguity is spatial relativity: data collected over shorter timeframes may occupy larger portions of the space where data are depicted and vice versa [48]. To avoid ambiguity, redundancy (the use of several data interactions that yield similar findings or interpretations) may be considered [49].

Biological processes seem to be unpredictable when classic methods are used [49–51]. However, methods that capture one-to-many/many-to-one relationships (which may change over time) may inform beyond the limitations of one-to-one, static methods [52–55].

2.7. From specialised (uni-disciplinary) to multi-, inter- and trans-disciplinary methods

The history of knowledge creation has followed a path that, recently, seemed to reach its limit. After the Humboldtian model of higher education requested published research as a fundamental requirement of new academic appointments, a rapid growth of specialties was observed, appearing to address. While attempts to bridge such gaps have fostered multidisciplinary approaches, such models do not necessarily share the concepts and language required to promote interactions across disciplines.

Inter-disciplinary approaches refer to those that not only foster a dialogue across disciplines but also solutions to ill-defined problems. When not only tangible solutions to complex problems are generated but also individuals not associated with any particular discipline participate, such methods are regarded as trans-disciplinary [57].

Therefore, one way to distinguish and evaluate scientific methods is investigating their temporal emphases. When they prioritize the past (prior knowledge), they tend to be unidisciplinary or specialized. Multi-disciplinary methods can also be past-oriented when the selection of disciplines invited to solve a problem is pre-established by an authority, as seen in many countries affected by COVID 19 [58].

In contrast, when problem-solving (future-oriented knowledge creation) is pursued, the method is viewed as inter- or trans-disciplinary. Unfortunately, academic institutions do not appear to strongly support such methods [59].

2.8. From data to information, knowledge, and beyond

The advent of the data deluge brought the concept that research methods could be reduced to data collection. In 2008, the intellectual process that also included research questions and explicit formulation of hypotheses seemed superfluous [20].

However, predicting the end of both theory and the scientific method has been challenged [17,36]. Data are now viewed as necessary but not sufficient to support informed decision-making. Instead, a process (the data-information-knowledge-wisdom or DIKW pyramid) is now regarded as essential to generate actionable knowledge [10,60]. While data, alone, lack meaning, structured data may inform, produce knowledge and, after being interpreted, generate wisdom (or understanding) that may be used in decision-making [61].

Yet, there is no consensus on the difference between data and information. While some authors have assigned the difference to functionality [61], other authors have claimed that data structuring is the central distinction between data and information [61]. Because information is, also, an ambiguous concept [62], structuring and functionality are not necessarily opposite to one another—they could be complementary, if not synonymous.

3. Section 2 – looking for concepts, working with variables: Contents and temporal stages of scientific methods

3.1. Connecting the abstract (unmeasurable) with the concrete (measurable) world

Methods are tentative idealizations of phenomena (abstract concepts), such as ‘health’, ‘immunity’, and ‘epidemic control.’ Ideally, researchers would like to measure such concepts. Because they are unmeasurable, investigators work with concrete (measurable) operations or variables, which may (or may not) relate to the concept(s) of interest. Fig. 1 shows how methods connect abstract with concrete elements.

The match (or lack of match) between concepts and operations has been a topic discussed for many decades. While operationalism has been abandoned as a philosophy of science, operationalizations subject to open-ended evaluations are still practiced [15]. For example, in control or prevention of epidemics, there is a potential mismatch between the concept ‘vaccines prevent infections’ and the operation expressed as ‘the

![Fig. 1. Four critical elements: concepts, variables, validations, and continuous revisions. Researchers never measure what they want to measure: they only measure a few variables that, at best, partially relate to the concept of interest. Therefore, validations and continuous evaluations are essential.](image-url)
concentration of neutralizing antibodies indicates whether protection has been achieved or not. Numerous biological factors may prevent protection even in vaccinated individuals [63].

3.2. The temporal stages -conceptualization, operationalization, validation, and evaluation

The connection between the abstract and the concrete world occurs through a process that includes four stages: conceptualization, operationalization, validation, and evaluation. Conceptualization is the process that identifies the concept of interest. When two or more of such concepts are selected, hierarchization takes place - a process in which concepts are ordered in a way such that the broadest one is located at the top of the hierarchy.

Once the conceptualization is concluded, operationalization follows. This is when the variable(s) most likely to capture the nature of the concept(s) of interest are selected, i.e., this is the stage in which data-generating variables are chosen.

For example, if a researcher wants to investigate a broad concept (e.g., ‘immunity’), subordinate concepts may be identified (e.g., ‘cell-mediated’ and ‘humoral immunity’). Then, concepts that facilitate double operations are chosen, such as ‘humoral immunity’ (an abstract concept) which may be estimated by ‘serum antibodies’ (measurable variables).

Operations are often numerical in nature and tend to offer mechanistic answers [64,65]. The operations designed by methods may include many elements, e.g., two or more (i) (subcellular to supra-cellular) biological levels, (ii) time points, and (iii) outcomes [66-68].

The conceptualization/operationalization process is followed by validation and evaluation. These two (sometimes indistinguishable) stages are motivated by the risks associated with selecting insufficient or irrelevant operations.

3.3. Validation

To validate, this question should be answered: are we measuring what we need to measure or what is easily measured but may be irrelevant? This implies a double ‘trip’ that includes (i) the transition between the abstract and the concrete world (from conceptualization to operationalization) and (ii) the reverse process, when the question mentioned above is addressed. When decision-makers initiate these processes, validations may be synonymous with evaluations (Fig. 2).

To estimate validity, at least five topics need to be considered: (i) the underlying theory, (ii) hypotheses, (iii) research design, (iv) empirical observations (data), and (v) revisions [69]. The underlying theory is expected to be consistent (‘coherent’) with the specific hypotheses being tested and also yield informative, explanatory, and/or usable data.

For example, if the theory is that ‘immunity, when effective, promotes survival’, one hypothesis could be that ‘multi-cellularity may be critical to foster survival.’ To test this hypothesis, a research design that explicitly produces multicellular variables would be needed, as previously shown [70]. Hence, future methodologies may be characterized by validations that, at least, pursue three goals: (i) predict outcomes; (ii) prescribe treatments, and (iii) understand or explain the underlying processes [71].

3.4. Data structuring and evaluation

Both validations and evaluations are likely to be influenced by the data. Earlier views on the scientific method did not consider the impact of data structure on the ‘DIKW’ pyramid [61]. When the original data are unstructured, they may lack functionality.

Gestalt theory-based pattern recognition fosters data structuring. Because it does not depend on numerical properties, Gestalt has solved problems that seemed intractable under mathematical approaches [72]. Pattern recognition is a top-down process, in which the overall shape (gestalt) is detected even when details are missing - a feature that results in novel information [40,72]. Fig. 3 is an example of a model meant to elicit data patterns.

3.5. Toward overall (inter-trans-disciplinary) knowledge validity

It is suggested that future scientific methodologies may need to explore their overall knowledge validity. That is so because what may seem to be adequate within the perspective of one field may reveal gaps when two or more perspectives are simultaneously considered.

One example of the previous concept is the relationship existing between public health and the economy: when one of these areas is neglected, both are deeply affected. In contrast, the benefits of population-level public health expenditure – unlike those of personal healthcare – tend to be long term and induce higher returns on the investment than alternatives [73]. It is now established that the economy requires public health [74]. COVID-19 has shown that both economic and health-related problems require rational decision-making [73].

The need for overall knowledge validity also applies to vaccine development. This field is not yet a science but a technology that is renovated empirically, i.e., it advances through trial-and-error processes [75]. While the COVID 19 pandemic has demonstrated its relevance, vaccinology has not been emphasized in biomedical curricula [76].

3.6. The twin needs: independent evaluations and biologically grounded methods

While no definition assures that a method is grounded in Biology, biologically grounded methods may be retrospectively validated by their findings. Methods that, analyzing the same data, extract more

Fig. 2. The double trip: back and forth between abstract concepts to concrete variables. To be valid, research should transition, twice, between the abstract and the concrete world.
information than alternatives and/or provide explanations consistent with Evolutionary theory may be assumed to be grounded in Biology [40].

Once data are structured following biological concepts -such as the fact that phagocytes predominate in the early inflammatory response while mononuclear cells show a higher presence in the recovery phase-, informative patterns tend to emerge. For instance, experimental and longitudinal data on infections induced in animals show a clear pattern (data circularity) in which the phagocyte/lymphocyte ratio and the neutrophil percentage reveal higher values in the early, post-challenge than the late/recovery stage, while the mononuclear cell/neutrophil ratio reaches its highest values in the recovery phase [Fig. 1 of reference # 47 (https://journals.plos.org/plosone/article?id=https://doi.org/10.1371/journal.pone.0053984)].

Biologically grounded and visually explicit methods may foster problem-solving [27,77]. For example, in epidemiology, the same data may be used to compare several theories (e.g., Network and Neighbor theories). Such approaches determine which theory is more explanatory [78]. Similarly, analyses of the same immunological data under under two (reductionist and non-reductionist) methods can detect the most informative [40]).

When one approach informs more or predicts better, research becomes indistinguishable from evaluation. However, to be effective, evaluations should be performed by independent systems [79].

3.7. Directionality matters: top-down and bottom up methods

To further elucidate whether a novel method is grounded in Biology, the directionality (bottom-up and/or top-down) of the process may influence on the detection of emergent biological properties. Top-down operations may uncover system-wide interactions. They are data-driven methods that attempt to solve ill-defined problems of unknown causation by focusing on spatial–temporal (toponomic or pattern recognition-based) features [80,81].

Top-down approaches can detect synergies, pleiotropies, and multicellularity –critical features of biological systems [82–85], and identify outcome-related patterns. Because Biomedicine is an indivisible system, top-down approaches help explore the interdependence of its elements [71,86].

However, validations may also require bottom-up operations. For instance, evaluations of health-related research utilize both top-down and bottom-procedures [87].

4. Conclusions

Because biomedical knowledge is complex and grows rapidly, biologically grounded scientific methods are needed and they should be frequently evaluated. To that end, the creation of novel, independent, and inter-/trans-disciplinary bodies of evaluators is proposed.

5. Author statement

Both authors have read and approved the final version of the manuscript and neither author has conflicts of interest to report. This manuscript has not been submitted to any other journal and it has not been funded by any institution—whether private or public—of any country.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] T.A.C. Raydon, How can science be well-ordered in times of crisis? Learning from the SARS CoV-2 pandemic, HPLS 42 (2020) S3, https://doi.org/10.1007/s40656-020-00348-5.
[2] J. Wise, Covid-19: Global response was too slow and leadership absent, report finds, BMJ 373 (2021).
[3] E.C. Sabino, I.F.R. Buss, M.P.S. Carvalho, C.A. Prete Jr, M.A.E. Cristi, N.A. Frajii, R.H.M. Pereira, K.V. Parag, et al., Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence (accessed: Feb 1, 2021), Lancet (2021), https://doi.org/10.1016/S0140-6736(21)00183-5.
[4] J.P.A. Ioannidis, S. Cripps, M.A. Tanner, Forecasting for COVID-19 has failed, Int. J. Forecast. (2020), https://doi.org/10.1016/j.ijforecast.2020.08.004.
