Synergistic use of data-based and mechanism-based modeling approaches

Vasilis Z. Marmarelis

In the history of science, the inductive (data-based) and the deductive (hypothesis-based) approaches have played a complementary and mutually beneficial role, whereby observations have led to the postulation of hypotheses that are subsequently tested by properly designed experiments, comprising an evolutionary cycle of hypothesis formulation and testing. In life sciences and medicine, the importance of discovering and quantifying the physiological mechanisms under normal and pathological conditions has given rise to mechanism-based modeling methods (e.g. compartmental modeling) which rely on experimental observations and properly formulated hypothesis testing. However, the intrinsic complexity of physiological systems has presented formidable challenges to this mechanism-based modeling approach and has motivated the development of data-based modeling approaches (typically input-output or stimulus-response formulations) that seek to capture the essential functional characteristics of the physiological system in a reliable manner consistent with the available data. Subsequent analysis of the obtained data-based models gives rise to hypothesis-based model forms that seek to encapsulate the relevant physiological mechanisms. The latter mechanism-based models are subsequently tested with data collected either from properly designed experiments or from spontaneous activity, leading either to validation or modification of the mechanism-based model in an evolutionary process where “model selection” is enabled by data-based modeling methods (i.e. the data-based model serves as “ground truth” for the validity/efficacy of a given mechanism-based model). Our experience over the last 30 years shows that this “virtuous cycle” of mechanism-based model development is enabled by the synergistic use of data-based and hypothesis-based models.

Several sources of complexity have been identified in the function of physiological systems. In addition to the cardinal issue of multiple interconnected variables that confound the analysis of a given system and the proper understanding of experimental data, the complicating presence of intrinsic dynamics, nonlinearities and nonstationarities of physiological systems are widely acknowledged as challenges in the effective and reliable modeling of these systems. For this reason, considerable effort has been dedicated in recent years to the development of data-based modeling methodologies that remain efficacious within this demanding context and facilitate the study of the relationships between the resulting data-based models and their hypothesis-based or mechanism-based counterparts. The proposed talk will summarize our findings to date and will present illustrative examples from neural and metabolic-endocrine systems where this synergistic approach has already yielded considerable benefits.

The issues of physiological complexity are germane to the studies of multi-scale modeling and strongly manifested in efforts to address the functional and structural complexity of living systems in a reliable and practicable manner within a hierarchical/orderly context of multiple scales of time and space. Although mechanism-based models remain the ultimate objective of multi-scale modeling, data-based methodologies using input-output model formulations can be very helpful in pursuing this goal, because of their applicability to arbitrary levels of systemic organization from molecular to cellular to multi-cellular to organ to multi-organ etc. This broad applicability hinges upon the availability of appropriate input-output data and effective methods of modeling/analysis within the constraints imposed by experimental limitations. This talk aims at stimulating our collective thinking on the synergistic use of data-based and mechanism-based modeling approaches that hold the promise for significant scientific advancement and translational opportunities leading to improved clinical practice.