Cell Systems

CellPress

Commentary

Numeracy 2.0—From analyzing data to evaluating biological insight

Alexander Hoffmann[1,*]
[1]Institute for Quantitative and Computational Biosciences, Department of Microbiology, Immunology, and Molecular Genetics; Molecular Biology Institute; Jonsson Comprehensive Cancer Center; and Broad Stem Cell Research Center, University of California, Los Angeles, Los Angeles, CA, USA
*Correspondence: ahoffmann@ucla.edu
https://doi.org/10.1016/j.cels.2026.101538

Biomedical research requires quantitative rigor, i.e., numeracy, a facility with numbers. The last decade has seen the broad adoption of statistical tools ("Numeracy 1.0"). To drive science forward, the expertise to quantitatively evaluate hypotheses and insights also needs to be broadly adopted ("Numeracy 2.0"). Systems biologists will be at the forefront of the transformation.

Numeracy 1.0: Measurements and data

The last decade(s) have witnessed a significant transformation of biomedical research in regard to the quantitative evaluation of experimental data. Two different causes are readily identified: first, prompted by a reproducibility crisis,[1,2] great effort has been made in improving the rigor of research by evaluating the statistical significance of measurements. These measurements—whether at the molecular, cellular, or tissue scale—provide a quantitative description of phenotype, and statistical evaluation determines the likelihood that the reported differences between two samples could have occurred by chance. All journals and study sections have guidelines so authors and reviewers pay attention to statistical significance when phenomena and phenotypes are described. The adoption of new standards has been so pronounced that papers that passed the peer review process even ten years ago would no longer satisfy expectations today. Further, while past research training may have included rudimentary familiarity with parametric tests, todays research training programs routinely include non-parametric tests, sampling approaches, mutual information, and regression methods (Numeracy 1.0).

The second impetus for the quantitative transformation has been the development of technology that produces large, high-dimensional datasets, increasingly at single-cell resolution. Research papers in any biomedical research area now routinely include such data, often generated by next-generation sequencing or imaging, which are then evaluated using sophisticated bioinformatic tools. Such tools address the challenges of data processing, dimensionality reduction, and clustering and may leverage external datasets or aggregated knowledge of interaction networks, pathways, or gene functions. Approaches powered by machine learning or deep learning (artificial intelligence) are adding further sophistication. Yet, while experimental biologists may learn how to use such tools (given well-documented tutorials), they rarely understand how they work or the underlying assumptions. There clearly remains an important role for computational biologists here, not only in the development of bioinformatic methods and their application but also in ensuring rigor to avoid a new reproducibility crisis.

Yet, while statistical methodologies (Numeracy 1,0) are providing more detailed and, in principle, more rigorous characterization of phenotypes, they do not directly address the question of how the phenotypes come about (Figure 1). Consulting knowledge bases can only offer candidate regulatory pathways that may potentially be involved; it does not allow for statements as to their requirement or sufficiency or what their fractional contribution may be.

Numeracy 2.0: Regulatory mechanisms

What underlies biological phenotypes and functions are dynamical systems of interacting cells or molecules that affect each other's abundance, location, or functional state; they generate dynamic responses to perturbations, provide homeostatic control that may be dysregulated in diseases, and mediate the effects of genetic variation and molecular therapeutics. Biology after all is about understanding regulatory mechanisms, not just collecting data about phenotypes.[3]

How does one apply quantitative rigor to insights about regulatory mechanism

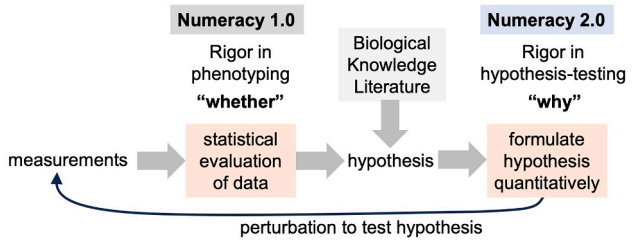

Figure 1. Distinguishing between different kinds of numeracies
Evaluating datasets with statistical tools (Numeracy 1.0) promotes rigor in evaluating experimental measurements, i.e., rigor in phenotyping. To evaluate biological insight quantitatively requires mechanistic models that formulate hypotheses quantitatively. This promotes rigor in hypothesis testing (Numeracy 2.0).

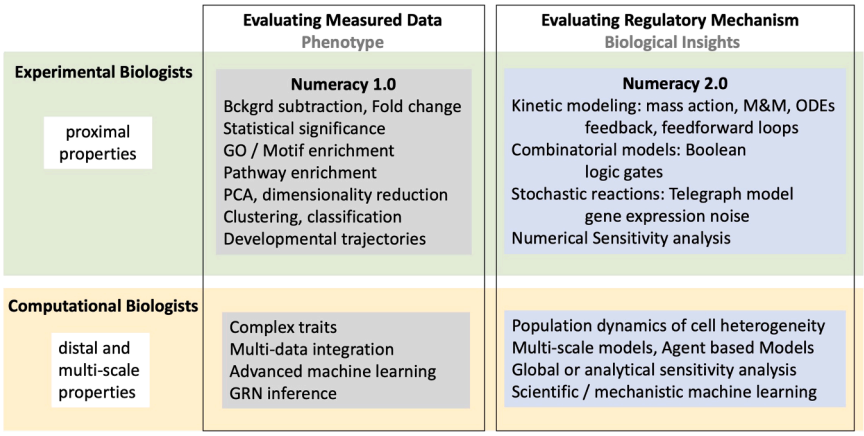

Figure 2. Leveraging the expertise of systems biologists for innovative research

As experimental biologists began to acquire statistical skillsets (Numeracy 1.0), computational biologists were free to focus on innovative research and methods development. By catalyzing Numeracy 2.0 among experimental biologists, computational biologists will be able to focus on complex problems that require complex methodologies and innovative algorithmic solutions.

underlying biological function? Of course, providing mechanistic insights is generally expected in research publications (except for "Resource" papers). Yet, while measurements are quantitatively evaluated for effect size and statistical significance, the contribution (effect size) of the hypothesized mechanism is rarely quantified, and a quantification of the confidence (statistical significance) associated with said mechanism is even less likely.

Consider for example the regulatory effect of transcription factors on stimulus-induced gene expression. The mRNA measurements are routinely evaluated quantitatively; thereby the effect size of the stimulus and the statistical significance are established. A loss-of-function mutation (e.g., by CRISPR) may establish the requirement of the transcription factor for this phenomenon of stimulus-induced mRNA abundance. But to what extent it actually mediates stimulus-induced transcription is unclear. That is because a fold change in mRNA abundance at a particular time point is not equivalent to changes in transcription rates, or the transcription factor may be required for licensing the induction but does not actually mediate it, and mRNA degradation may also be indrectly affected. Similar questions remain when regulatory mechanisms are reported for cellular behavior such as cell death/survival, proliferation, chemotaxis, etc. or higher-level functions such as pathogen resistance, tumor elimination, or autoimmune phenomena.

So how are regulatory mechanisms quantitatively evaluated? What is the equivalent of a Student's *t* test to evaluate a hypothesis? Instead of a model of intersecting distributions used for the Student's *t* test, a mathematical model that captures the proposed mechanism and links it to the phenomena and measurements is required. Such a model can quantify the regulatory effect and temporal relationships of the proposed mechanism and determine how robustly it fits available knowledge and data. Importantly, it makes underlying simplifying assumptions explicit; thereby it lends itself to comparisons of different model conceptions encoding different regulatory mechanisms in accounting for the available data and prior knowledge. It thereby begins to evaluate the confidence we have in the hypothesized regulatory mechanism (Figure 1). With new datasets, the sufficiency of the previously published regulatory mechanism can be determined. Thus, gains in knowledge are made explicit.

It is remarkable that while basic calculus generally precedes basic statistics in high school or college mathematics, in bio-scientific practice basic statistics for data evaluation is an assumed skillset for all experimentalists, but the use of differential equations is largely confined to the field of systems biology. Indeed, even the smallest mathematical equation will trigger the transfer of a research paper or grant application to a specialist review panel or journal or at least the inclusion of a mathematical modeler, who may not have any expertise in the biological subject. While the inclusion of statistical tests for the quantitative evaluation of measurements and data is a requirement for all biological research, the inclusion of mathematical models for the quantitative evaluation of regulatory mechanisms is not. In that, the inclusion of a quantitative evaluation of the proposed mechanism pigeonholes the research as "specialized" systems biology.

These current scientific research and review practices reveal a lack of familiarity with a basic, college-level mathematical understanding of calculus, and associated concepts and notations, within the research community. It represents a numeracy gap that is distinct from that which was addressed for evaluating the statistical significance of data comparisons. Numeracy 2.0 is a call to further improve the quantitative literacy of the biomedical research community to improve the rigor of biomedical research.

Systems biologists are key catalysts

Biological systems are complex, consisting of numerous cellular agents each with regulatory networks comprising numerous molecules. Regulatory processes have direct targets and direct effects that occur in the short term and indirect or downstream effects that may occur longer term or on a different scale. We may distinguish them as "proximal systems properties" or "distal systems properties," respectively (Figure 2). Size-scale and timescale separation is a key strategy for identifying proximal regulatory mechanisms within a complex biological process. Here, I would like to suggest that the focus of systems biologists is shifting from proximal to distal systems properties, with the expectation that experimental biologists should step up their Numeracy 2.0 to quantitatively evaluate proximal systems properties.

For instance, within the unit of the cell, proximal systems properties of signaling and gene regulatory networks have been the focus of avid systems biology research. How do the molecular reactions within intracellular networks control the activation of transcription factors or the expression of immune response genes? Based on my own experience of the last 25 years, mathematical models

are key to interpreting and understanding the regulatory basis of dose-response relationships, the numerous functional effects of negative feedback,[4] the generation of stimulus-specific transcription factor dynamics,[5] and how these dynamics are interpreted via fold change detection by feedforward loops[6] or slow chromatin dynamics.[7,8] Dynamical systems have hysteresis and may have bistability, which in the regulatory networks of immune cells manifest themselves as mechanisms for memory and timekeeping.

On another scale, how rates of cellular proliferation or differentiation account for cell population dynamics over time has identified cell stages of proliferative expansion[9] and identified enhanced or slowed differentiation as mechanisms of reduced or enhanced mature cell production. As proliferative processes are exponential, the relationship between mechanism and phenomena is inherently non-linear.

The repertoire of mathematical conceptions for such regulatory systems is mature, not unlike the repertoire of available statistical tests, and the software available to implement these is well established. As such, the quantitative characterization of proximal emergent systems properties should be within the realm of experimental biologists who have subject expertise and who are experts at generating appropriate datasets to test these models.

Distal systems properties, however, require more complex quantitative modeling strategies that may involve algorithmic innovations to dissect their regulatory mechanisms. Moving between scales, from molecular networks to cell populations, requires accounting for cell-to-cell heterogeneity, distinguishing between measurement imprecision, or molecular noise that is extrinsic of or intrinsic to the biological system under study.[10–12] In such projects, systems biologists may ask: how do genes and reactions within intracellular molecular networks control cell population dynamics and/or biological functions at the physiological scale? How do metabolic networks within cells interface with each other through nutrient availability and sharing of intermediate metabolites within the tumor micro-environment? Or how does epigenetic state control of lymphocytes interface with their genetic selection in the thymus and germinal centers during immune development and vaccine or pathogen responses? The latter exemplifies particularly well the complexities that arise when selection favors outliers of heterogenous cell distribution and couples that with an exponential growth process. This breaks the ergodicity assumption of general mathematical modeling. While explicit agent-based models are a reliable yet computationally expensive and sometimes intractable approach, more succinct multi-scale modeling requires expert training in the art of the craft of dynamical systems modeling.

In sum, if experimentalists are being trained in the statistical evaluation of their datasets, one should also expect that they can quantitatively evaluate the regulatory mechanisms that control proximal systems properties. While some level of competence has been achieved in the former (Numeracy 1.0), our attention should focus on achieving competence in the latter (Numeracy 2.0), as both should be expected to be a component of every biomedical research project. More complex tasks of high-dimensional data integration of numerous sources through innovative dimensional reduction, harmonization, and data-driven modeling approaches, as well as the analysis of the regulatory mechanisms that control distal systems emergent properties, will likely remain the realm of expert computational systems biologists, bio-data scientists, or mathematical biologists.

Next steps for Numeracy 2.0

What will it take to improve numeracy in regard to evaluating biological insight and regulatory mechanisms? As systems biologists, we can look back at the playbook that established Numeracy 1.0, i.e., quantitative rigor in evaluating measurements. It involved (1) reviewer guidelines by journal and study sections so that authors describe the statistical tests used and data are deposited to allow for reproduction by third parties, (2) courses and workshops in statistical methods in undergraduate and graduate programs, and (3) biostatistics cores and consulting services in major research centers. In essence, the inclusion of statistics was not considered to be "interdisciplinary" but was incorporated into the discipline of biomedical sciences.

Similarly, we may advocate for making the quantitative evaluation of biological insight, hypothesis, or regulatory mechanisms as the norm for studies that report them. Rather than seeing such research that involves mathematical modeling as interdisciplinary, we hope to see it as an integral component of biomedical sciences going forward.

To advance Numeracy 2.0, we may advocate the following:

1. Reviewer guidelines by journals and study sections so that authors describe the approaches to quantify regulatory relationships, making assumptions explicit and evaluating alternate hypotheses to establish confidence and deposition of models
2. Courses and workshops in mechanistic modeling and dynamical systems, using toolboxes in popular platforms of R and python, the same used as for statistical evaluation of the data
3. Collaborative support services to consult; train; and ensure rigor in model building, implementation, and evaluation

At present, quantitative descriptions of regulatory mechanisms in experimental biomedical studies are rare. There are no generally accepted guidelines or standards to establish rigor in reporting regulatory mechanisms. For the next phase of advancing biomedical scientific practice, we need to expand numeracy from describing phenotypes to describing regulatory relationships.

The systems biology community, i.e., the readership of *Cell Systems*, is key to advancing this transformation that focuses on the numeracy of regulatory insights. Systems biologists can help catalyze and broaden expertise, help establish standards, and also push back when being asked to evaluate simple models of proximal emergent properties in order to educate reviewers, editors, and study section coordinators. In turn, this will free up the systems biology community to focus on research areas of more distal emergent properties that are innovative and potentially multi-scale or address biological heterogeneity. Numeracy 2.0 spells out an exciting goal for the next 10 years!

ACKNOWLEDGMENTS

I acknowledge formative discussions with my UCLA colleagues at the QCBio Institute, especially Eric Deeds and Roy Wollman, as well as Harinder Singh (U Pittsburgh); critical reading by Francisco Lopes (Federal University Rio de Janeiro); and my own ongoing training by my numerous students and postdoctoral fellows. I acknowledge current NIH funding: R01AI173214, R01AI185026, U19AI172713, and P01AI120944.

DECLARATION OF INTERESTS

The author declares no competing interests.

REFERENCES

1. Ioannidis, J.P.A. (2005). Why Most Published Research Findings Are False. PLoS Med. *2*, e124. https://doi.org/10.1371/journal.pmed.0020124.

2. Baker, M. (2016). 1,500 Scientists Lift the Lid on Reproducibility. Nature *533*, 452–454. https://doi.org/10.1038/533452a.

3. Brenner, S. (2003). Nobel Lecture: Nature's Gift to Science. Biosci. Rep. *23*, 225–237.

4. Basak, S., Behar, M., and Hoffmann, A. (2012). Lessons from Mathematically Modeling the NF-κB Pathway. Immunol. Rev. *246*, 221–238. https://doi.org/10.1111/j.1600-065X.2011.01092.x.

5. Adelaja, A., Taylor, B., Sheu, K.M., Liu, Y., Luecke, S., and Hoffmann, A. (2021). Six Distinct NFκB Signaling Codons Convey Discrete Information to Distinguish Stimuli and Enable Appropriate Macrophage Responses. Immunity *54*, 916–930.e7. https://doi.org/10.1016/j.immuni.2021.04.011.

6. Lee, R.E.C., Walker, S.R., Savery, K., Frank, D.A., and Gaudet, S. (2014). Fold Change of Nuclear NF-κB Determines TNF-Induced Transcription in Single Cells. Mol. Cell *53*, 867–879. https://doi.org/10.1016/j.molcel.2014.01.026.

7. Sen, S., Cheng, Z., Sheu, K.M., Chen, Y.H., and Hoffmann, A. (2020). Gene Regulatory Strategies That Decode the Duration of NFκB Dynamics Contribute to LPS- versus TNF-Specific Gene Expression. Cell Syst. *10*, 169–182.e5. https://doi.org/10.1016/j.cels.2019.12.004.

8. Cheng, Q.J., Ohta, S., Sheu, K.M., Spreafico, R., Adelaja, A., Taylor, B., and Hoffmann, A. (2021). NF-κB Dynamics Determine the Stimulus Specificity of Epigenomic Reprogramming in Macrophages. Science (New York, N.Y.) *372*, 1349–1353. https://doi.org/10.1126/science.abc0269.

9. Höfer, T., Busch, K., Klapproth, K., and Rodewald, H.-R. (2016). Fate Mapping and Quantitation of Hematopoiesis In Vivo. Annu. Rev. Immunol. *34*, 449–478. https://doi.org/10.1146/annurev-immunol-032414-112019.

10. Mitchell, S., Roy, K., Zangle, T.A., and Hoffmann, A. (2018). Nongenetic Origins of Cell-to-Cell Variability in B Lymphocyte Proliferation. Proc. Natl. Acad. Sci. USA *115*, E2888–E2897. https://doi.org/10.1073/pnas.1715639115.

11. Singh, A., Chia, J.J., Rao, D.S., and Hoffmann, A. (2025). Population Dynamics Modeling Reveals That Myeloid Bias Involves Both HSC Differentiation and Progenitor Proliferation Biases. Blood *145*, 1293–1308. https://doi.org/10.1182/blood.2024025598.

12. Guo, X., Adelaja, A., Singh, A., Wollman, R., and Hoffmann, A. (2025). Modeling Heterogeneous Signaling Dynamics of Macrophages Reveals Principles of Information Transmission in Stimulus Responses. Nat. Commun. *16*, 5986. https://doi.org/10.1038/s41467-025-60901-3.