

Editorial Commentary: Machine Learning in Orthopaedics: Venturing Into the Valley of Despair



Ian J. Wellington, M.D., and
Mark P. Cote, P.T., D.P.T., M.S.C.T.R., Associate Editor, Statistics

Abstract: Machine learning, a subset of artificial intelligence, has become increasingly common in the analysis of orthopaedic data. The resources needed to utilize machine-learning approaches for data analysis have become increasingly accessible to researchers, contributing to a recent influx of research using these techniques. As machine learning becomes increasingly available, misapplication owing to a lack of competence becomes more common. Sensationalized titles, misused vernacular, and a failure to fully vet machine learning–derived algorithms are just a few issues that warrant attention. As the orthopaedic community’s knowledge on this topic grows, the flaws in our understanding of this field will likely become apparent, allowing for rectification and ultimately improvement of how machine learning is utilized in research.

See related article on page 2761

The article “Meaningless Applications and Misguided Methodologies in Artificial Intelligence–Related Orthopaedic Research Propagates Hype Over Hope” by Ramkumar, Pang, Polisetty, Helm, and Karnuta¹ takes a critical look at artificial intelligence (AI) in the realm of orthopaedics. AI is one of the most fascinating—yet possibly least understood—areas of orthopaedic research. Machine learning (ML), a subset of AI, has grown in popularity as a method of data analysis in the literature.^{2–7} Despite the growth of ML, Ramkumar et al. paint a grim picture of the current state of ML-related orthopaedic research. Sensationalized titles, misused vernacular, and a failure to fully vet ML-derived algorithms are just a few issues that warrant attention.

As technology advances, ML programming becomes increasingly user-friendly, allowing the previously uninitiated to take a crack at these techniques. The Dunning-Kruger effect is a psychological bias that correlates confidence and competence for a specific topic (Fig 1).⁸ It suggests that persons with relatively

little competence in a subject may have high levels of confidence because of unawareness of the breadth of knowledge contained within that field. This point of high confidence and low competence is colloquially referred to as the “Peak of Mount Stupid.” As competence increases, one becomes aware of how much there is to know and confidence drops precipitously into the “Valley of Despair.” Over time, competence and confidence gradually improve, progressing along the “Slope of Enlightenment” until one reaches the “Plateau of Sustainability,” indicating true mastery of a subject. The hype-fueled cycle described by Ramkumar et al.¹ has turned out a growing number of ML-related studies replete with errors in basic fundamentals, raising concern that we may be collectively sitting at the peak of the Dunning-Kruger curve, unaware of the true complexities associated with ML and ultimately AI in general.

So here we sit, arguably atop the peak. As the research community becomes increasingly familiar with ML over time, we will begin our descent into the Valley of Despair, where we will become aware of the deficiencies in our understanding of ML methodologies. However, as the gaps in knowledge become apparent, we will begin to fill them, producing more meaningful and clinically relevant research. As Ramkumar et al.¹ point out, there is a strong trend to publish commentaries and reviews on ML in orthopaedics. Although the irony that this very editorial is a commentary on a

University of Connecticut (I.J.W.).

The authors report the following potential conflicts of interest or sources of funding: M.P.C. is a paid committee member of AANA, outside the submitted work, and receives financial compensation from AANA. Full ICMJE author disclosure forms are available for this article online, as [supplementary material](#).

© 2022 by the Arthroscopy Association of North America
0749-8063/22666/\$36.00

<https://doi.org/10.1016/j.arthro.2022.05.010>

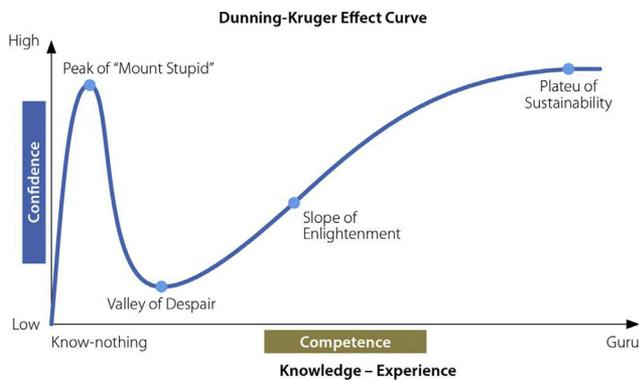


Fig 1. Dunning-Kruger effect curve.

commentary about ML is apparent, the prevalence of these sorts of articles does serve a purpose. Although they do not progress our orthopaedic knowledge, they do explain some of the complexities and pitfalls of ML, educating the community and progressing researchers' overall understanding of this topic. As this knowledge progresses, we continue progressing along the Dunning-Kruger curve, beyond the Valley of Despair, to a true understanding of ML and its capabilities.

Acknowledgment

The authors acknowledge Dr. Geneva Hargis for providing the illustration for this article.

References

1. Ramkumar P, Pang M, Polisetty T, Helm JM, Karnuta J. Meaningless applications and misguided methodologies in artificial intelligence-related orthopaedic research propagates hype over hope. *Arthroscopy* 2022;38:2761-2766.
2. Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. *JAMA* 2017;318:517-518.
3. Cabitza F, Locoro A, Banfi G. Machine learning in orthopedics: A literature review. *Front Bioeng Biotechnol* 2018;6:75.
4. Kunze KN, Polce EM, Rasio J, Nho SJ. Machine learning algorithms predict clinically significant improvements in satisfaction after hip arthroscopy. *Arthroscopy* 2021;37:1143-1151.
5. Lu Y, Lavoie-Gagne O, Forlenza EM, et al. Duration of care and operative time are the primary cost drivers after ambulatory hip arthroscopy: A machine learning analysis. *Arthroscopy* 2022;38:2204-2216.e3.
6. Grazal CF, Anderson AB, Booth GJ, Geiger PG, Forsberg JA, Balazs GC. A machine-learning algorithm to predict the likelihood of prolonged opioid use following arthroscopic hip surgery. *Arthroscopy* 2022;38:839-847.
7. Cote MP, Lubowitz JH, Brand JC, Rossi MJ. Artificial intelligence, machine learning, and medicine: A little background goes a long way toward understanding. *Arthroscopy* 2021;37:1699-1702.
8. Kruger J, Dunning D. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol* 1999;77:1121.