

Viewpoint

Is Artificial Intelligence Better Than Human Clinicians in Predicting Patient Outcomes?

Joon Lee^{1,2,3}, PhD

¹Data Intelligence for Health Lab, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

²Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

³Department of Cardiac Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

Corresponding Author:

Joon Lee, PhD

Data Intelligence for Health Lab

Cumming School of Medicine

University of Calgary

3280 Hospital Dr NW

TRW 5E17

Calgary, AB, T2N 4Z6

Canada

Phone: 1 403 220 2968

Email: joonwu.lee@ucalgary.ca

Abstract

In contrast with medical imaging diagnostics powered by artificial intelligence (AI), in which deep learning has led to breakthroughs in recent years, patient outcome prediction poses an inherently challenging problem because it focuses on events that have not yet occurred. Interestingly, the performance of machine learning–based patient outcome prediction models has rarely been compared with that of human clinicians in the literature. Human intuition and insight may be sources of underused predictive information that AI will not be able to identify in electronic data. Both human and AI predictions should be investigated together with the aim of achieving a human-AI symbiosis that synergistically and complementarily combines AI with the predictive abilities of clinicians.

(*J Med Internet Res* 2020;22(8):e19918) doi: [10.2196/19918](https://doi.org/10.2196/19918)

KEYWORDS

patient outcome prediction; artificial intelligence; machine learning; human-generated predictions; human-AI symbiosis

Absence of Human-Generated Predictions in Patient Outcome Research

In recent years, there has been a proliferation of patient outcome prediction research that applies machine learning (ML) and artificial intelligence (AI) to electronic health records (EHRs) and other clinical and administrative health data. The central premises are that 1) complex health data contains predictive information that ML can effectively extract and transform into a predictive algorithm and 2) accurate prediction of patient outcomes can facilitate early, preventative intervention and more efficient health care resource allocation through identification of high-risk patients. For example, predicting which intensive care unit patients are likely to develop sepsis can prompt early initiation of fluid resuscitation, vasopressor therapy, or antibiotics, which can reduce damage from insufficient organ perfusion [1,2]. Although AI has been

enormously successful in medical imaging diagnostics, where the medical condition of interest is already present or absent in the images (eg, diagnosis of diabetic retinopathy [3] and classification of skin lesions [4]), patient outcome prediction poses an inherent challenge of predicting events that have *not yet* occurred (eg, mortality, length of stay, and readmission) [5]. This challenge is common to both AI and human clinicians.

Interestingly, while human and AI predictions are often directly compared in medical imaging research [6-8], patient outcome prediction studies tend to focus only on ML and seldom investigate human predictions. This is corroborated by a number of systematic reviews and meta-analyses, which target only ML methods [9-14] or empirical methods [15-19]. This gap in the literature is coherent across a wide range of medical specialties and diseases, including trauma [9], cancer [11], neurosurgery [10], depression [12], acute gastrointestinal bleeding [13], sepsis [14], acute liver failure [15], ischemic stroke [16], thermal injury

[17], and cardiovascular disease [18,19]. The absence of human predictions appears to be a recent trend, as older literature prior to the current widespread use of modern ML and EHRs includes more comparisons of human and AI predictions [20-23].

There are several possible reasons why human performance is more frequently studied in medical imaging than in patient outcome prediction. First, radiologists are trained to analyze, interpret, and classify images, whereas most other medical specialists are not trained to directly predict patient outcomes. While accurate prognostic information can certainly be helpful in any medical specialty, it is usually generated by empirical risk scoring systems such as the Framingham Risk Score [24] or Acute Physiology and Chronic Health Evaluation (APACHE) [25] rather than by human clinicians. Second, human predictions in medical imaging are readily available from routine clinical practice or can be generated systematically by trained radiologists. Conversely, it is rare for clinicians in other medical specialties to record patient outcome predictions that *they* generate on a regular basis. Third, the implicit assumption is that humans cannot accurately predict patient outcomes because analysis of complex, high-dimensional clinical data may be required; moreover, recall bias is rampant in the human mind.

Humans and AI Should Work as a Team

However, there is no reason to rule out the possibility that human clinicians can outperform AI in patient outcome prediction, at least in some clinical scenarios. While AI can only access information that can be recorded in the form of electronic data, human clinicians interact face-to-face with their patients and have access to both clinical and contextual information. The qualitative information collected via clinicians' five senses can be critical in patient outcome prediction; however, this information is mostly absent in EHRs, if it is possible to record it at all. Although some qualitative observations can be recorded in EHRs as free-text notes, such as nursing notes, these data are logged in a limited, inconsistent fashion. Human intuition and insight may well be the most underused resources in patient outcome prediction.

While the performance of ML-based patient outcome prediction models appears impressive on paper, the most accurately predicted cases tend to be "easy" cases where the likely outcomes are already obvious to human clinicians [26]. This further supports the hypothesis that human clinicians perform well in patient outcome prediction.

On the other hand, AI easily outperforms humans in processing, analyzing, and finding patterns in complex, high-dimensional data [27]. As demonstrated by IBM Watson [28] and AlphaGo [29], the memory, attention, and information processing abilities of AI vastly exceed the capabilities of human cognition [30]. This AI advantage is crucial for extracting and using data-driven insights from big data [31]; it is also key to the recent successful breakthroughs in ML, particularly in deep learning [32], in a number of problem domains, including medical imaging [33]. In addition, AI does not suffer from fatigue [34] or cognitive biases (eg, recall bias) [35] as humans do. However, even if AI outperforms human clinicians in patient outcome prediction, human performance represents a more meaningful benchmark

that puts AI performance in better perspective. Understanding the superiority of AI in comparison with humans can facilitate adoption of AI technology in real patient care.

The bottom line is that both AI and humans can make unique contributions to patient outcome prediction, and they should help each other to maximize predictive performance. Patient outcome prediction research should aim for human-AI symbiosis, where the respective predictive abilities of AI and human clinicians are combined in a synergistic and complementary way [36]. Given the challenging nature of patient outcome prediction, creating an AI to act alone without human help will simply lead to suboptimal predictive performance because even state-of-the-art ML technology cannot leverage information that is not present in the data [26].

Another way for AI and humans to work together is via the human-in-the-loop model, where humans directly inform machines on how to learn from the data at hand by providing guidance based on human intuition and knowledge. The term "interactive machine learning" [37] was coined to describe this paradigm; it encompasses more well-known branches of ML, such as active learning, where humans select which data points should be labelled. This human-in-the-loop approach can greatly reduce the computational complexity of some ML problems; for example, it has shown promising results in protein folding [38]. Moreover, in the field of human-computer interaction, the human-in-the-loop concept has been studied in the context of vehicle control [39], security [40,41], and decision-making [40,42]. Knowledge from these application areas can potentially inform the design of human-AI symbiosis in patient outcome prediction.

AI and human prediction performance may vary across different types of patients. Complex patterns in data can be more predictive than human intuition in certain patient subgroups, and the opposite may be true in other subpopulations. An investigation of how AI and human predictions can be optimally combined for different types of patients could directly contribute to advancing precision medicine. A better understanding of the respective predictive powers of AI and humans in various clinical scenarios can also help increase human trust in AI (eg, "For this type of patient, I need to trust AI more because most predictive information is buried in the complex data"). This can facilitate evidence-based adoption of AI technology.

For human clinicians to completely trust AI, it is necessary to understand *why* an algorithm arrives at a given conclusion; this requires transparency, traceability, and causality. The active field of explainable AI has been producing useful methods, such as SHapley Additive exPlanations (SHAP) [43], that can help explain how ML models work at an algorithmic level (this explanation is almost always based on correlation rather than causation); however, human clinicians ultimately want to elevate this algorithmic explainability to a model that is understandable by humans with sufficient causal understanding, also known as causability [44]. Therefore, mapping explainability to causability will be key in achieving true human-AI symbiosis.

One major roadblock to the proposed human-AI symbiosis is the need to collect a large number of human predictions in a variety of clinical scenarios, which is labor-intensive and adds

to clinicians' workloads. Seamlessly integrated electronic prediction collection platforms (eg, embedded in a multi-center EHR system) can minimize this burden and enable large-scale prediction collection.

From Patient Outcome Prediction to Real Impact

Once predictive performance is optimized via human-AI symbiosis, the next important step is to formulate clinical guidelines so that the predictive information is actionable. This is a crucial step, as accurate predictions alone will not lead to

any real impact; rather, the combination of accurate predictions and appropriate interventions by clinicians will have a greater effect [5,26].

The ultimate goal of patient outcome prediction is to improve patient outcomes and decrease health care costs through early intervention and efficient use of health care resources. To prove that this goal has been met, we will need to perform randomized clinical trials of AI-driven patient care [45], such as that conducted by Wijnberge and colleagues [46]. In addition to simply comparing AI with human work alone, these randomized clinical trials should investigate a promising third species: human-AI symbiosis.

Acknowledgments

The author would like to thank the University of Calgary for institutional support.

Conflicts of Interest

None declared.

References

1. Desautels T, Calvert J, Hoffman J, Jay M, Kerem Y, Shieh L, et al. Prediction of Sepsis in the Intensive Care Unit With Minimal Electronic Health Record Data: A Machine Learning Approach. *JMIR Med Inform* 2016 Sep 30;4(3):e28 [FREE Full text] [doi: [10.2196/medinform.5909](https://doi.org/10.2196/medinform.5909)] [Medline: [27694098](https://pubmed.ncbi.nlm.nih.gov/27694098/)]
2. Lee J, Mark RG. An investigation of patterns in hemodynamic data indicative of impending hypotension in intensive care. *Biomed Eng Online* 2010;9(1):62. [doi: [10.1186/1475-925x-9-62](https://doi.org/10.1186/1475-925x-9-62)]
3. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016 Dec 13;316(22):2402-2410. [doi: [10.1001/jama.2016.17216](https://doi.org/10.1001/jama.2016.17216)] [Medline: [27898976](https://pubmed.ncbi.nlm.nih.gov/27898976/)]
4. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017 Jan 25;542(7639):115-118. [doi: [10.1038/nature21056](https://doi.org/10.1038/nature21056)]
5. Lindsell CJ, Stead WW, Johnson KB. Action-Informed Artificial Intelligence-Matching the Algorithm to the Problem. *JAMA* 2020 May 01;323(21):2141. [doi: [10.1001/jama.2020.5035](https://doi.org/10.1001/jama.2020.5035)] [Medline: [32356878](https://pubmed.ncbi.nlm.nih.gov/32356878/)]
6. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digit Health* 2019 Oct;1(6):e271-e297. [doi: [10.1016/s2589-7500\(19\)30123-2](https://doi.org/10.1016/s2589-7500(19)30123-2)]
7. Emblem KE, Nedregaard B, Hald JK, Nome T, Due-Tønnessen P, Bjørnerud A. Automatic glioma characterization from dynamic susceptibility contrast imaging: brain tumor segmentation using knowledge-based fuzzy clustering. *J Magn Reson Imaging* 2009 Jul;30(1):1-10. [doi: [10.1002/jmri.21815](https://doi.org/10.1002/jmri.21815)] [Medline: [19557840](https://pubmed.ncbi.nlm.nih.gov/19557840/)]
8. Emblem KE, Pinho MC, Zöllner FG, Due-Tønnessen P, Hald JK, Schad LR, et al. A generic support vector machine model for preoperative glioma survival associations. *Radiology* 2015 Apr;275(1):228-234. [doi: [10.1148/radiol.14140770](https://doi.org/10.1148/radiol.14140770)] [Medline: [25486589](https://pubmed.ncbi.nlm.nih.gov/25486589/)]
9. Liu NT, Salinas J. Machine Learning for Predicting Outcomes in Trauma. *Shock* 2017 Nov;48(5):504-510. [doi: [10.1097/SHK.0000000000000898](https://doi.org/10.1097/SHK.0000000000000898)] [Medline: [28498299](https://pubmed.ncbi.nlm.nih.gov/28498299/)]
10. Senders JT, Staples PC, Karhade AV, Zaki MM, Gormley WB, Broekman ML, et al. Machine Learning and Neurosurgical Outcome Prediction: A Systematic Review. *World Neurosurg* 2018 Jan;109:476-486.e1. [doi: [10.1016/j.wneu.2017.09.149](https://doi.org/10.1016/j.wneu.2017.09.149)] [Medline: [28986230](https://pubmed.ncbi.nlm.nih.gov/28986230/)]
11. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J* 2015;13:8-17 [FREE Full text] [doi: [10.1016/j.csbj.2014.11.005](https://doi.org/10.1016/j.csbj.2014.11.005)] [Medline: [25750696](https://pubmed.ncbi.nlm.nih.gov/25750696/)]
12. Lee Y, Ragguett R, Mansur R, Boutilier J, Rosenblat J, Trevizol A, et al. Applications of machine learning algorithms to predict therapeutic outcomes in depression: A meta-analysis and systematic review. *J Affect Disord* 2018 Dec 01;241:519-532. [doi: [10.1016/j.jad.2018.08.073](https://doi.org/10.1016/j.jad.2018.08.073)] [Medline: [30153635](https://pubmed.ncbi.nlm.nih.gov/30153635/)]
13. Shung D, Simonov M, Gentry M, Au B, Laine L. Machine Learning to Predict Outcomes in Patients with Acute Gastrointestinal Bleeding: A Systematic Review. *Dig Dis Sci* 2019 Aug;64(8):2078-2087. [doi: [10.1007/s10620-019-05645-z](https://doi.org/10.1007/s10620-019-05645-z)] [Medline: [31055722](https://pubmed.ncbi.nlm.nih.gov/31055722/)]

14. Islam M, Nasrin T, Walther B, Wu C, Yang H, Li Y. Prediction of sepsis patients using machine learning approach: A meta-analysis. *Comput Methods Programs Biomed* 2019 Mar;170:1-9. [doi: [10.1016/j.cmpb.2018.12.027](https://doi.org/10.1016/j.cmpb.2018.12.027)] [Medline: [30712598](https://pubmed.ncbi.nlm.nih.gov/30712598/)]
15. Wlodzimirow KA, Eslami S, Chamuleau RAFM, Nieuwoudt M, Abu-Hanna A. Prediction of poor outcome in patients with acute liver failure-systematic review of prediction models. *PLoS One* 2012;7(12):e50952 [FREE Full text] [doi: [10.1371/journal.pone.0050952](https://doi.org/10.1371/journal.pone.0050952)] [Medline: [23272081](https://pubmed.ncbi.nlm.nih.gov/23272081/)]
16. Fahey M, Crayton E, Wolfe C, Douiri A. Clinical prediction models for mortality and functional outcome following ischemic stroke: A systematic review and meta-analysis. *PLoS One* 2018;13(1):e0185402 [FREE Full text] [doi: [10.1371/journal.pone.0185402](https://doi.org/10.1371/journal.pone.0185402)] [Medline: [29377923](https://pubmed.ncbi.nlm.nih.gov/29377923/)]
17. Hussain A, Choukairi F, Dunn K. Predicting survival in thermal injury: a systematic review of methodology of composite prediction models. *Burns* 2013 Aug;39(5):835-850. [doi: [10.1016/j.burns.2012.12.010](https://doi.org/10.1016/j.burns.2012.12.010)] [Medline: [23384617](https://pubmed.ncbi.nlm.nih.gov/23384617/)]
18. van Dieren S, Beulens J, Kengne A, Peelen L, Rutten G, Woodward M, et al. Prediction models for the risk of cardiovascular disease in patients with type 2 diabetes: a systematic review. *Heart* 2012;360-369. [doi: [10.1136/heartjnl-2011-300734](https://doi.org/10.1136/heartjnl-2011-300734)]
19. Damen JAAG, Hooft L, Schuit E, Debray TPA, Collins GS, Tzoulaki I, et al. Prediction models for cardiovascular disease risk in the general population: systematic review. *BMJ* 2016 May 16;353:i2416 [FREE Full text] [doi: [10.1136/bmj.i2416](https://doi.org/10.1136/bmj.i2416)] [Medline: [27184143](https://pubmed.ncbi.nlm.nih.gov/27184143/)]
20. Grove W, Zald D, Lebow B, Snitz B, Nelson C. Clinical versus mechanical prediction: A meta-analysis. *Psychol Assess* 2000;12(1):19-30. [doi: [10.1037/1040-3590.12.1.19](https://doi.org/10.1037/1040-3590.12.1.19)]
21. Gardner W, Lidz CW, Mulvey EP, Shaw EC. Clinical versus actuarial predictions of violence in patients with mental illnesses. *J Consult Clin Psychol* 1996;64(3):602-609. [doi: [10.1037/0022-006x.64.3.602](https://doi.org/10.1037/0022-006x.64.3.602)]
22. Marchese MC. Clinical versus actuarial prediction: a review of the literature. *Percept Mot Skills* 1992 Oct;75(2):583-594. [Medline: [1408625](https://pubmed.ncbi.nlm.nih.gov/1408625/)]
23. Bandiera G, Stiell IG, Wells GA, Clement C, De Maio V, Vandemheen KL, et al. The Canadian C-Spine rule performs better than unstructured physician judgment. *Ann Emerg Med* 2003 Sep;42(3):395-402. [doi: [10.1016/s0196-0644\(03\)00422-0](https://doi.org/10.1016/s0196-0644(03)00422-0)]
24. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation* 2008 Feb 12;117(6):743-753. [doi: [10.1161/CIRCULATIONAHA.107.699579](https://doi.org/10.1161/CIRCULATIONAHA.107.699579)] [Medline: [18212285](https://pubmed.ncbi.nlm.nih.gov/18212285/)]
25. Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: Hospital mortality assessment for today's critically ill patients*. *Crit Care Med* 2006;34(5):1297-1310. [doi: [10.1097/01.ccm.0000215112.84523.f0](https://doi.org/10.1097/01.ccm.0000215112.84523.f0)]
26. Chen JH, Asch SM. Machine Learning and Prediction in Medicine - Beyond the Peak of Inflated Expectations. *N Engl J Med* 2017 Jun 29;376(26):2507-2509 [FREE Full text] [doi: [10.1056/NEJMp1702071](https://doi.org/10.1056/NEJMp1702071)] [Medline: [28657867](https://pubmed.ncbi.nlm.nih.gov/28657867/)]
27. Jarrahi MH. Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Bus Horiz* 2018 Jul;61(4):577-586. [doi: [10.1016/j.bushor.2018.03.007](https://doi.org/10.1016/j.bushor.2018.03.007)]
28. Ferrucci D, Brown E, Chu-Carroll J, Fan J, Gondek D, Kalyanpur A, et al. Building Watson: An Overview of the DeepQA Project. *AIMag* 2010 Jul 28;31(3):59-79. [doi: [10.1609/aimag.v31i3.2303](https://doi.org/10.1609/aimag.v31i3.2303)]
29. Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, et al. Mastering the game of Go with deep neural networks and tree search. *Nature* 2016 Jan 28;529(7587):484-489. [doi: [10.1038/nature16961](https://doi.org/10.1038/nature16961)] [Medline: [26819042](https://pubmed.ncbi.nlm.nih.gov/26819042/)]
30. Grigsby SS. Artificial intelligence for advanced human-machine symbiosis. In: *Augmented Cognition: Intelligent Technologies*. AC 2018. Lecture Notes in Computer Science. 2018 Presented at: International Conference on Augmented Cognition; July 15-20, 2018; Las Vegas, NV p. 255-266. [doi: [10.1007/978-3-319-91470-1_22](https://doi.org/10.1007/978-3-319-91470-1_22)]
31. L'Heureux A, Grolinger K, Elyamany H, Capretz M. Machine Learning With Big Data: Challenges and Approaches. *IEEE Access* 2017;5:7776-7797. [doi: [10.1109/access.2017.2696365](https://doi.org/10.1109/access.2017.2696365)]
32. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015 May 28;521(7553):436-444. [doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539)] [Medline: [26017442](https://pubmed.ncbi.nlm.nih.gov/26017442/)]
33. Martín Noguero T, Paulano-Godino F, Martín-Valdivia MT, Menias CO, Luna A. Strengths, Weaknesses, Opportunities, and Threats Analysis of Artificial Intelligence and Machine Learning Applications in Radiology. *J Am Coll Radiol* 2019 Sep;16(9 Pt B):1239-1247. [doi: [10.1016/j.jacr.2019.05.047](https://doi.org/10.1016/j.jacr.2019.05.047)] [Medline: [31492401](https://pubmed.ncbi.nlm.nih.gov/31492401/)]
34. Hockey G, Wiethoff M. Cognitive fatigue in complex decision-making. *Adv Space Biol Med* 1993:139-150. [doi: [10.1016/S1569-2574\(08\)60101-X](https://doi.org/10.1016/S1569-2574(08)60101-X)]
35. Dawson NV, Arkes HR. Systematic errors in medical decision making. *J Gen Intern Med* 1987 May;2(3):183-187. [doi: [10.1007/bf02596149](https://doi.org/10.1007/bf02596149)]
36. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019 Jan 7;25(1):44-56. [doi: [10.1038/s41591-018-0300-7](https://doi.org/10.1038/s41591-018-0300-7)] [Medline: [30617339](https://pubmed.ncbi.nlm.nih.gov/30617339/)]
37. Holzinger A. Interactive machine learning for health informatics: when do we need the human-in-the-loop? *Brain Inform* 2016 Jun;3(2):119-131 [FREE Full text] [doi: [10.1007/s40708-016-0042-6](https://doi.org/10.1007/s40708-016-0042-6)] [Medline: [27747607](https://pubmed.ncbi.nlm.nih.gov/27747607/)]
38. Cooper S, Khatib F, Treuille A, Barbero J, Lee J, Beenen M, et al. Predicting protein structures with a multiplayer online game. *Nature* 2010 Aug 05;466(7307):756-760 [FREE Full text] [doi: [10.1038/nature09304](https://doi.org/10.1038/nature09304)] [Medline: [20686574](https://pubmed.ncbi.nlm.nih.gov/20686574/)]

39. Driggs-Campbell K, Shia V, Bajcsy R. Improved driver modeling for human-in-the-loop vehicular control. In: 2015 IEEE International Conference on Robotics and Automation (ICRA). 2015 Presented at: IEEE International Conference on Robotics and Automation (ICRA); 2015; Seattle, WA p. 1654-1661. [doi: [10.1109/icra.2015.7139410](https://doi.org/10.1109/icra.2015.7139410)]
40. Schumann MA, Drusinsky D, Michael JB, Wijesekera D. Modeling Human-in-the-Loop Security Analysis and Decision-Making Processes. IIEEE Trans. Software Eng 2014 Feb;40(2):154-166. [doi: [10.1109/tse.2014.2302433](https://doi.org/10.1109/tse.2014.2302433)]
41. Cranor LF. A Framework for Reasoning About the Human in the Loop. usenix.org. 2008. URL: https://www.usenix.org/legacy/event/upsec/tech/full_papers/cranor/cranor.pdf [accessed 2020-07-24]
42. Subramania HS, Khare VR. Pattern classification driven enhancements for human-in-the-loop decision support systems. Decision Support Systems 2011 Jan;50(2):460-468. [doi: [10.1016/j.dss.2010.11.003](https://doi.org/10.1016/j.dss.2010.11.003)]
43. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From Local Explanations to Global Understanding with Explainable AI for Trees. Nat Mach Intell 2020 Jan;2(1):56-67 [FREE Full text] [doi: [10.1038/s42256-019-0138-9](https://doi.org/10.1038/s42256-019-0138-9)] [Medline: [32607472](https://pubmed.ncbi.nlm.nih.gov/32607472/)]
44. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. Wiley Interdiscip Rev Data Min Knowl Discov 2019;9(4):e1312 [FREE Full text] [doi: [10.1002/widm.1312](https://doi.org/10.1002/widm.1312)] [Medline: [32089788](https://pubmed.ncbi.nlm.nih.gov/32089788/)]
45. Angus DC. Randomized Clinical Trials of Artificial Intelligence. JAMA 2020 Feb 17;323(11):1043-1045. [doi: [10.1001/jama.2020.1039](https://doi.org/10.1001/jama.2020.1039)] [Medline: [32065828](https://pubmed.ncbi.nlm.nih.gov/32065828/)]
46. Wijnberge M, Geerts B, Hol L, Lemmers N, Mulder M, Berge P, et al. Effect of a Machine Learning-Derived Early Warning System for Intraoperative Hypotension vs Standard Care on Depth and Duration of Intraoperative Hypotension During Elective Noncardiac Surgery: The HYPE Randomized Clinical Trial. JAMA 2020 Feb 17:1052-1060. [doi: [10.1001/jama.2020.0592](https://doi.org/10.1001/jama.2020.0592)] [Medline: [32065827](https://pubmed.ncbi.nlm.nih.gov/32065827/)]

Abbreviations

- AI:** artificial intelligence
APACHE: Acute Physiology And Chronic Health Evaluation
EHR: electronic health record
ML: machine learning
SHAP: SHapley Additive exPlanations

Edited by G Eysenbach; submitted 06.05.20; peer-reviewed by A Holzinger, Z Ge; comments to author 12.06.20; revised version received 24.06.20; accepted 25.06.20; published 26.08.20

Please cite as:

Lee J

Is Artificial Intelligence Better Than Human Clinicians in Predicting Patient Outcomes?

J Med Internet Res 2020;22(8):e19918

URL: <http://www.jmir.org/2020/8/e19918/>

doi: [10.2196/19918](https://doi.org/10.2196/19918)

PMID: [32845249](https://pubmed.ncbi.nlm.nih.gov/32845249/)

©Joon Lee. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 26.08.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.