DOI: 10.1111/jocs.16844

COMMENTARY

JOURNAL OF CARDIAC SURGERY WILEY

Will artificial intelligence help us in predicting outcomes in cardiac surgery?

Carlos A. Mestres MD, PhD, FETCS[1] | Eduard Quintana MD, PhD, FETCS[1] | Daniel Pereda MD, PhD[2]

[1]Department of Cardiothoracic Surgery and the Robert WM Frater Cardiovascular Research Centre, The University of the Free State, Bloemfontein, South Africa

[2]Department of Cardiovascular Surgery, Hospital Clinic, The University of Barcelona, Barcelona, Spain

Correspondence: Carlos A. Mestres, MD, PhD, FETCS, Department of Cardiothoracic Surgery, Faculty of Health Sciences, The University of the Free State, Internal Box G 32, Bloemfontein 9300, South Africa.
Email: carlosamestres@gmail.com

KEYWORDS
artificial intelligence, cardiac surgery, machine learning, outcomes

Computers changed our lives for good and for bad quite a while ago. People in the kind of seniority side may agree that the Commodore 64, also known as the C64 or the CBM 64, an 8-bit home computer introduced in January 1982 by Commodore International, was the home computer that allowed many of us to be introduced in regular daily computing. At some point, the Commodore 64 was the best-selling home computer.[1] Commodore International, an American computer and electronics manufacturer, was instrumental in the development and global acceptance of the personal computer industry in the 1970s and 1980s. It was operational between 1958 and 1994, when it ceased operations after the company announced voluntary bankruptcy.[2] Other computer platforms replaced it. Today we are almost at the mercy of computers as our daily life is almost controlled by them in different ways.

Artificial intelligence (AI) is the human-like behavior displayed by a machine or system.[3] For the dummy, like those of us who signed this brief commentary, computers are aiming to mimic human behavior using as many data as possible from other examples that behave similarly. The goal is that machines should be able to work on the extreme side of efficiency by analyzing large, vast amounts of data in the shortest period of time.[3] Historically, one may say that the seminal contribution of the late Alan Turing in 1936 paved the way for the development of what computing is today.[4] This seems to be his most important contribution ever that was, of course, highlighted by many with the occasion of his centenary.[5,6]

With the understanding that history is useful for us to understand what happened in the past that we did not witness, AI is an old and very powerful tool in any field where large data are continuously generated. Medicine is a good example. Machine learning (ML), a form of AI, is rapidly making inroads and is going to be a part of the future in Medicine. It is currently used in different applications in healthcare. ML shows some potential ability to predict specific complications such as pancreatic fistula after pancreaticoduodenectomy.[7] Deep learning software as a subset of ML seems to confirm around 80% sensitivity and specificity in detecting intracranial aneurysms on digital subtraction angiography as per a recent feasibility study.[8] In the field of imaging in cardiology, studies suggest that ML may play a role in predicting follow-up outcomes and, generally speaking improving patient care.[9] Finally, AI seems to produce subjective image quality in the evaluation of pulmonary nodules with the proviso that there may be eventual false-positive findings.[10] Pooling all together, AI will likely and hopefully positively impact healthcare.

Having said that, in this issue of the Journal, Penny-Dimri et al. present our readership with an interesting systematic review and meta-analysis on how ML may be a helpful tool for the prediction of outcomes in cardiac surgery.[11] Cardiac surgery is a specialty that has produced substantial amount of data resulting in different tools to predict outcomes before an operation is performed. The European System for Cardiac Operative Risk Evaluation (EuroSCORE additive, logistic and II)[12] and the Society of Thoracic Surgeons predicted risk of mortality (STS-PROM)[13] have been in use for a couple of decades

and have been incorporated in the daily surgical practice due to their reliability in predicting outcomes with a given set of patient-related and operation-related variables. Of interest is that Penny-Dimri et al.,[11] digging into a bunch of studies identifying 51 of value, did not find differences in outcome prediction, namely 30-day mortality and in-hospital mortality, between "traditional" statistical models and ML methods. In other words, ML has no greater discriminative power than logistic regression.

The main question here is why ML methods are not superior to the traditional statistical tools that we currently manage in cardiac surgery. Authors speculate with the slow adoption of such a technology, which is probably true in Medicine in comparison with other branches of science or industry. Other problems identified in this meta-analysis are the type of studies identified. Authors included observational cross-sectional, case–control, and cohort studies comparing any ML algorithm to any reference standard or a different ML algorithm. The studies were either retrospective or prospective with no randomized studies incorporated. This is simply what was available at the time of their analysis.

Looking at different studies previously mentioned here, authors are careful at the moment in issuing strong conclusions and recommendations.[7–10] Although feasibility seems not to be an issue, there are still problems with interpretation and the possibility of identifying false-positive findings. Outcomes in cardiac surgery are a complex topic. As rightly stated by the authors, we may still miss cases of mortality outside of the measurement boundary and a good example are interhospital transfers. Above all interpretations is the fact that AI/ML need huge amount of data; this is a constant across the studies scrutinized by the authors. Low numbers are the norm in cardiac surgery in comparison with other fields. If one understands that, according to social media management platforms, people globally watch over one billion hours of YouTube videos every day —we speak about two billion people—,[14] it is easy to understand that the low numbers, as well discussed by Penny-Dimri et al.[11] are a fundamental limitation of ML methods in predicting outcomes in cardiac surgery now. Of course, using YouTube as a reference or comparator is a silly approach for different reasons; the first and most important is that YouTube is not a reliable source of medical-related information as appropriately documented by Osman et al.[15]

Penny-Dimri et al.[11] have to be commended as they have looked at the currently available information on ML methods and outcome prediction. Screening over 2000 articles is not an easy task as time is gold; however, and assuming the intrinsic limitations of meta-analyses, their conclusion is strong, ML methods are not superior to currently statistical methods for outcome prediction in cardiac surgery. For the time being, then, stick to what we do on a daily basis and wait for further improvement in methodology due to still doubtful utility.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Carlos A. Mestres http://orcid.org/0000-0001-8148-9044
Eduard Quintana http://orcid.org/0000-0002-5386-7463
Daniel Pereda http://orcid.org/0000-0002-0526-8722

REFERENCES

1. Griggs B.The commodore 64, that '80s computer icon, lives again. *CNN*. May 9, 2011.
2. Burgess J.Adios amiga and commodore: from a bang to a whimper, PC maker closes its doors. *The Washington Post*. May 9, 1994.
3. hpe.com/ae/en/what-is/artificial-intelligence.html?jumpid
4. Turing AM. On computable numbers with an application to the *Entscheidungsproblem*. *Proc Lond Math Soc*. 1937;s2-42:230-265.
5. Dyson G. Turing centenary: the dawn of computing. *Nature*. 2012;482(7386):459-460. doi:10.1038/482459a
6. Brenner S. Turing centenary: life's code script. *Nature*. 2012;482 (7386):461. doi:10.1038/482461a
7. Kambakamba P, Mannil M, Herrera PE, et al. The potential of machine learning to predict postoperative pancreatic fistula based on preoperative, non-contrast-enhanced CT: a proof-of-principle study. *Surgery*. 2020;167:448-454. doi:10.1016/j.surg.2019.09.019
8. Hainc N, Mannil M, Anagnostakou V, et al. Deep learning based detection of intracranial aneurysms on digital subtraction angiography: a feasibility study. *Neuroradiol J*. 2020;33:311-317. doi:10.1177/1971400920937647
9. Mannil M, Eberhard M, von Spiczak J, Heindel W, Alkadhi H, Baessler B. Artificial intelligence and texture analysis in cardiac imaging. *Curr Cardiol Rep*. 2020;22:131. doi:10.1007/s11886-020-01402-1
10. Jungblut L, Blüthgen C, Polacin M, et al. First performance evaluation of an artificial intelligence-based computer-aided detection system for pulmonary nodule evaluation in dual-source photon-counting detector CT at different low-dose levels. *Invest Radiol*. 2022;57:108-114. doi:10.1097/RLI.0000000000000814
11. Penny-Dimri JC, Bergmeir C, Perry L, Hayes L, Bellomo R, Smith JA. Machine learning to predict adverse outcomes after cardiac surgery: a systematic review and meta-analysis. *J Thorac Cardiovasc Surg* . 2022;163(6):2075-2087.e9.
12. Cromhout PF, Thygesen LC, Moons P, Nashef S, Damgaard S, Berg SK. Social and emotional factors as predictors of poor outcomes following cardiac surgery. *Interact Cardiovasc Thorac Surg*. 2022;34:193-200. doi:10.1093/icvts/ivab261
13. Jacobs JP, Shahian DM, O'Brien SM, Badhwar V. Assessing the performance of risk models with discrimination and calibration. *Ann Thorac Surg*. 2022;(22): 00800-1. doi:10.1016/j.athoracsur.2022.05.032
14. https://blog.hootsuite.com/how-the-youtube-algorithm-works/#How_does_the_YouTube_algorithm_work_in_2022
15. Osman W, Mohamed F, Elhassan M, Shoufan A. Is YouTube a reliable source of health-related information? A systematic review. *BMC Med Educ*. 2022;22:382. doi:10.1186/s12909-022-03446-z