

EDITORIAL



Clinical management of sepsis can be improved by artificial intelligence: yes

Matthieu Komorowski^{1,2*}

© 2019 Springer-Verlag GmbH Germany, part of Springer Nature

The management of sepsis is a highly complex, multifaceted challenge that remains the realm of highly skilled and trained human experts. But as medical applications of artificial intelligence continue to pour in, it is becoming obvious that some of these decisions could soon be left to machines that could be dubbed “intelligent”, improving clinical practice and patient outcomes [1]. Indeed, most of the tasks involved in the clinical management of sepsis (early recognition, selection of antibiotic therapy, haemodynamic optimisation, etc.) could be individually performed or optimised by dedicated algorithms.

Most of what we call “artificial intelligence” is in fact machine learning—a set of computer tools intended to generate new knowledge from data [1]. Machine learning includes three categories of techniques: supervised (which uses labelled data to build a prediction model, for example for prognostication), unsupervised (which discovers patterns in data and generates clusters of subjects that share common characteristics) and reinforcement learning (where a sequential decision process is modelled and optimised).

Below, I have selected a few significant applications that I consider the most likely to land in the clinical environment in the near future, either because of their robustness or their potential.

Notable applications of artificial intelligence in sepsis

First of all, it appears possible for automated algorithms to identify patients at risk of having sepsis, either in real time (“sepsis detection”) or in advance (“sepsis

prediction”) [2]. This can be achieved with a range of supervised learning algorithms trained on a dataset containing negative and positive instances of sepsis. For example, a model based on gradient tree boosting showed good accuracy for predicting sepsis and septic shock using only vital signs several hours before onset [3]. Even simpler rule-based algorithms are capable of highlighting at-risk patients, for example by detecting end-organ damage and the non-specific Systemic Inflammatory Response Syndrome [4], quick Sequential Organ Failure Assessment (qSOFA) or SOFA scores [5]. Remarkably, the small study by Shimabukuro and colleagues was one of the few to demonstrate improved outcomes with an algorithm from a randomised trial, which is an important lesson: successful deployment and acceptance by clinicians are much more important than the intrinsic complexity of an algorithm [4].

Next, unsupervised learning is a collection of tools used to identify subgroups of patients with sepsis, arguably a highly generic syndrome. Various teams have proposed to apply clustering algorithms to separate patients based on their clinical, biological or omic data [6, 7]. While most of this research remains exploratory and hypothesis generating, it now appears possible to envision practical use. Antcliffe used transcriptomic data to show that a subgroup of patients had poorer outcomes when given steroids, which may find clinical application in the future [7].

The cornerstone of the haemodynamic management of sepsis remains to this day the optimisation of the circulating blood volume with intravenous fluids and/or vasopressors. Correct timing and dosing of these medications are highly challenging, and it appears that reinforcement learning could help with this difficult task [8]. In reinforcement learning, the disease process of septic patients is modelled into multiple health states, then the decisions taken by human clinicians are analysed and their value

*Correspondence: m.komorowski14@imperial.ac.uk

² Intensive Care Unit, Charing Cross Hospital, Fulham Palace Road, London W6 8RF, UK

Full author information is available at the end of the article

quantified [8]. Finally, decisions more likely to lead to improved organ function and/or survival are identified. A different approach could use complex causal inference methods to identify the best time for starting vasopressors, by quantifying the relative risk of death for different timing options [9].

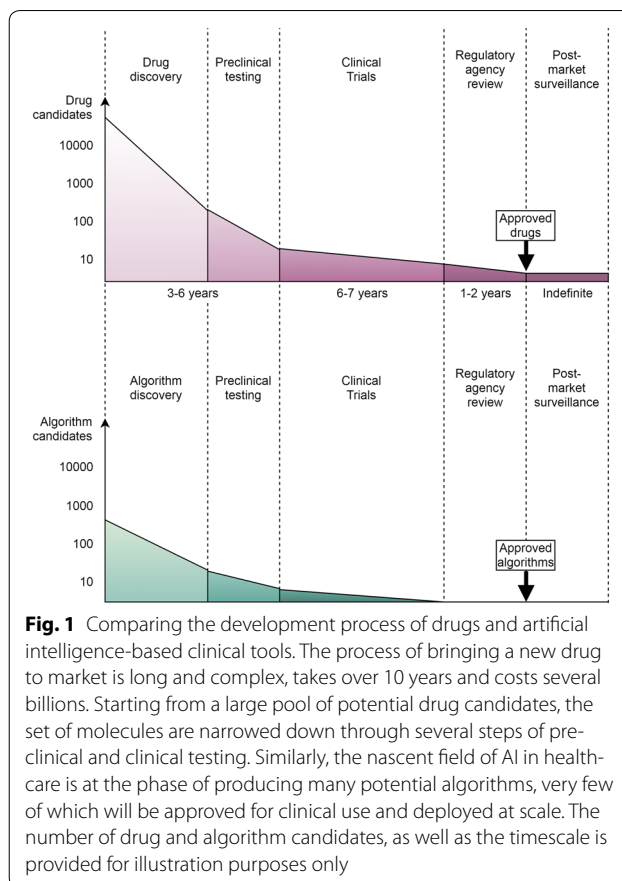
Assessing the value of new decisions in a purely retrospective fashion is challenging: when the algorithm and clinicians disagree, it is difficult to reliably estimate what would have happened to the patient. Retrospective validation relies on expert assessment of the model behaviour, sensitivity analyses to test its robustness and a type of statistical methods called “off-policy evaluation” to estimate the value of the algorithm’s decisions, all of which have limitations [8, 10].

What next? The route to clinical deployment

The examples listed above have some potential for improving the outcomes of septic patients, but prospective validation in the clinical environment is widely lacking (the “byte to bedside” gap), so evidence of efficacy and safety in the real world remain scarce [1, 11]. The path to regulation (e.g. C.E. marking) is long and complex and beyond the scope of this paper, but regulatory bodies are getting up to speed and several products are now cleared for clinical use [1]. In parallel, the question of the legal framework is deferred by leaving the responsibility of the decision to clinicians.

Let us outline what the route to market could look like. First, an algorithm needs to be externally validated, which poses the non-trivial challenges of accessing and compatibilising different datasets. Then, prospective testing could be conducted in “silent” mode, by off-duty clinicians without informing patient care. The next step would be a randomised controlled trial comparing physicians alone to physicians assisted by the algorithm. Besides clinical benefit, the inventors will have to demonstrate that their tool provides interpretability, confidence intervals in its estimates and can detect deviations from a required behaviour. Provided it does, this should warrant sufficient evidence for approval by the regulatory bodies.

A parallel can be drawn between the traditional drug development pathway by pharmaceutical companies and the burgeoning field of “medical artificial intelligence” (Fig. 1). The process of bringing a new medication to market is long and difficult, with most ideas being discarded along the way. Similarly, we can expect most published algorithms to never impact clinical care, and indeed very few systems are currently approved for clinical use [1]. What we need, if nothing else, is more algorithms, more teams testing them in a systematic, safe and controlled fashion, and more support from funding and regulatory agencies.



Once deployed, these systems will be available to clinicians as real-time bedside tools, for example, accelerating sepsis recognition or suggesting a dose range of intravenous fluids. If the evidence of their superiority is strong enough, we can even anticipate that the use of these systems might become enforced, similar to what exists for example in the aviation industry with automated landing systems [12].

To conclude, the future is bright for the field of artificial intelligence in medicine and applications in sepsis are nearing clinical deployment. If we manage to produce and test enough candidate systems, then success will become unavoidable and we will see the day where patients with sepsis are optimally managed by tandems of algorithms and human clinicians working hand in hand.

Author details

¹ Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London SW7 2AZ, UK. ² Intensive Care Unit, Charing Cross Hospital, Fulham Palace Road, London W6 8RF, UK.

Compliance with ethical standards

Conflicts of interest

The author declares that he has no conflict of interest.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 16 October 2019 Accepted: 9 December 2019

Published online: 13 December 2019

References

1. Topol EJ (2019) High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 25:44. <https://doi.org/10.1038/s41591-018-0300-7>
2. Islam MM, Nasrin T, Walther BA et al (2019) Prediction of sepsis patients using machine learning approach: a meta-analysis. *Comput Methods Programs Biomed* 170:1–9. <https://doi.org/10.1016/j.cmpb.2018.12.027>
3. Mao Q, Jay M, Hoffman JL et al (2018) Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department, general ward and ICU. *BMJ Open* 8:e017833. <https://doi.org/10.1136/bmjopen-2017-017833>
4. Shimabukuro DW, Barton CW, Feldman MD et al (2017) Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial. *BMJ Open Respir Res* 4:e000234. <https://doi.org/10.1136/bmjresp-2017-000234>
5. Donnelly JP, Safford MM, Shapiro NI et al (2017) Application of the Third International Consensus Definitions for Sepsis (Sepsis-3) Classification: a retrospective population-based cohort study. *Lancet Infect Dis* 17:661–670. [https://doi.org/10.1016/S1473-3099\(17\)30117-2](https://doi.org/10.1016/S1473-3099(17)30117-2)
6. Seymour CW, Kennedy JN, Wang S et al (2019) Derivation, validation, and potential treatment implications of novel clinical phenotypes for sepsis. *JAMA* 321:2003–2017. <https://doi.org/10.1001/jama.2019.5791>
7. Antcliffe DB, Burnham KL, Al-Beidh F et al (2018) Transcriptomic signatures in sepsis and a differential response to steroids: from the VANISH randomized trial. *Am J Respir Crit Care Med*. <https://doi.org/10.1164/rccm.201807-1419OC>
8. Komorowski M, Celi LA, Badawi O et al (2018) The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med* 24:1716. <https://doi.org/10.1038/s41591-018-0213-5>
9. Waschka A, Komorowski M, Hubbard A, Pirracchio R (2019) Optimal timing to vasopressor initiation in septic shock: revisiting SEPSIS-3 guidelines using machine learning. *Intensive Care Med Exp* 7(Suppl 3):001685
10. Gottesman O, Johansson F, Meier J, et al (2018) Evaluating reinforcement learning algorithms in observational health settings. arXiv:1805.12298 [cs, stat]
11. Coiera E (2019) The last mile: where artificial intelligence meets reality. *J Med Internet Res* 21:e16323. <https://doi.org/10.2196/16323>
12. International Civil Aviation Organization (2013) ICAO 9365—manual of all-weather operations. International Civil Aviation Organization, Montreal