Discover Health Systems

Perspective

Evidence pyramid and artificial intelligence: a metamorphosis of clinical research

Valentina Bellini[1] · Enrico Ori[1] · Federico Coccolini[2] · Elena Bignami[1]

Received: 19 September 2023 / Accepted: 16 October 2023
Published online: 21 November 2023

Abstract

We are in the era of disruption. New technologies are all around us and they change our professional and non-professional life very quickly. Evidence-based medicine is changing accordingly. We propose a new version of the evidence pyramid that evolves through the changes of our novel perspective. New technologies are manipulating the pyramid, revealing us unexpected dimensions, and shaping it into a more complex form. From a traditional two-dimensional pyramid, we move on to a three-dimensional one, where the third dimension is represented by the research efforts to go up to the next step. Legal, ethical, educational and cost-effectiveness issues are some of the major current barriers to manage and remove from accessing a secure and daily routine use of these intelligent tools. An additional element, represented by the volume of the step, highlighting the impact in clinical practice of each level of the pyramid, is added creating the fourth new version of evidence pyramid.

Keywords Artificial intelligence · Medical research · Evidence-based medicine · AI implementation · Clinical research · Evidence pyramid

1 Introduction

This is the era of disruption: more changes in different areas of our life, from trade to art [1], have taken place and the other will come [2]. Artificial Intelligence (AI), Internet-of-Medical Things (IoMT) and the other widespread technologies are trying to enter our daily life. Clinical research is also evolving too. Traditional analysis methods are not sufficient to value and analyze the large number of datasets available today [3]. In fact, data that can be stored have undergone a gradual but decisive change in their amount and format. From a list of intermittent static standardized data, today we have at our disposal real time and multi-source data interconnected together.

It is something like cinema: from single black-and-white frames we moved on to continuous footage enriched with colors, sounds and music up to 3D vision and multi-sensory experiences. As in the film industry, technological progression has made it possible to store and manage an increasing amount of data in medicine as well. Alongside traditional healthcare data, we have information from various other sources, such as personal devices, social media, biobanks and so on. There is a transition from an approach focused on intermittent data to one based on an increasing amount of structured and unstructured real time data [4]. We are moving towards real personalized medicine; considering all the characteristics of the subjects, thanks to new technologies, it is possible to developing models for prevention, diagnosis

✉ Elena Bignami, elenagiovanna.bignami@unipr.it | [1]Anesthesiology, Critical Care and Pain Medicine Division, Department of Medicine and Surgery, University of Parma, Viale Gramsci 14, 43126 Parma, Italy. [2]Department of General, Emergency and Trauma Surgery, Pisa University Hospital, 56124 Pisa, Italy.

and treatment extremely tailor-made. It is no coincidence that these algorithms often perform as well as, or even better than, clinicians when it comes to answering specific questions [5].

In parallel with this technological revolution, the concept of Evidence-Based Medicine, which has maintained medical practice for decades, could not fail to change. If EBM represents an emblematic sign of the Third Industrial Revolution [6], AI and Maching Learning are those of the Fourth Industrial Revolution [7]. Subbiah V. recently presented a brilliant overview of how medical research is changing in recent years, highlighting how disruptive technologies will be among the key factors of this evolution [8]. Revolution, in our opinion, which involves a gradual but profound change in the way of doing research and where AI and other new technologies will be allies of healthcare professionals with the patients' health as common purpose. However, there is a current underuse of these techniques both in clinical practice and research. Probably one of the main problems is that currently many of us feel invested by these technologies rather than being able to ride them. And while this might be perceived as acceptable in everyday life, for example AI that manages our emails or the use of intelligent voice assistants as Apple's Siri® or Amazon's Alexa®, most of us are not willing to employ technologies that do not understand and which there is not yet a significant amount in the evidence in their work. Currently there seems to be a kind of tug-of-war, between those who would like to adopt these technologies daily, and those who still do not have faith in them and think we are too distant from their use in clinical and research practice. We are probably in a transitional phase; our mission is to find the best way for human intelligence, artificial intelligence and other new technologies to create a hybrid healthcare model, to ensure that future generations can make the most of this medical technology revolution.

The aim of this paper is to highlight the ongoing modification of the pyramid of evidence and to suggest a possible future evolution of research, and subsequently clinical practice, with the introduction and implementation of AI and new technologies.

2 AI integration in pyramid of evidence: the construction of a 2D-AI-EBM

Medical practice heavily depends on the best evidence available in research; pyramid of evidence is well known, being used traditionally in order to describe studies' strength in terms of reliability, trustworthiness and reduction of bias (Fig. 1a).

Evidence gains more and more strength as epidemiological accuracy and statistical precision increase, ascending step by step until reaching the top [9].

Some necessary and important changes in the concept of the evidence pyramid are taking place with the introduction of AI in healthcare [10] (Fig. 1b).

For example, there have been great opportunities with respect to a case report with the advent of in silico studies. With these studies, in fact, we can observe and evaluate the behavior and effects of certain molecules on an organism created specifically through the application of AI. This gives the opportunity to carry out research with broad clinical implications in a short time, with few resources, with infinite possibilities to change the starting point.

For example, Williams et al. in their silico study investigates the use of vancomycin in Intensive Care Unit by comparing a newly developed dosing guideline with product-information-based dosing, using a pharmacokinetic model based on a seriously ill patient population. In ninety-six simulations, the authors measured predefined therapeutic, subtherapeutic, and toxicity pharmacokinetic-pharmacodynamic (PK-PD) targets. They found that guideline dosage was slightly more effective than standard dosing, as per product information, in achieving PK-PD exposure associated with an increased likelihood of effectiveness. Additionally, the first one significantly reduced the risk of subtherapeutic exposure. The possibility of surpassing toxicity thresholds, however, was higher with the guidelines, and additional research is recommended to enhance dosing precision and sensitivity [11].

With these various new research methods, it is also feasible to comprehensively examine the clinical hazard experienced by a patient with specific traits and illnesses. Of course, all of this allows for a more personalized approach than before. All this could change clinical-diagnostic behaviors in some areas where, for example, expert opinions or case reports were the basis for therapeutic choice. An instance of this is Burlacu et al.'s publication, where they emphasize that antithrombotic treatments for advanced chronic kidney disease (CKD) patients with cardiovascular comorbidities can be addressed by utilizing Machine Learning (LM) algorithms to create a new bleeding risk algorithm. This provides more coherent data than singular case reports [6].

The application of these methods will not only be able to stratify risk and select optimal treatment pathways for patients, but also improve treatment response and Quality of Life (QoL). In fact, predictive algorithms have been

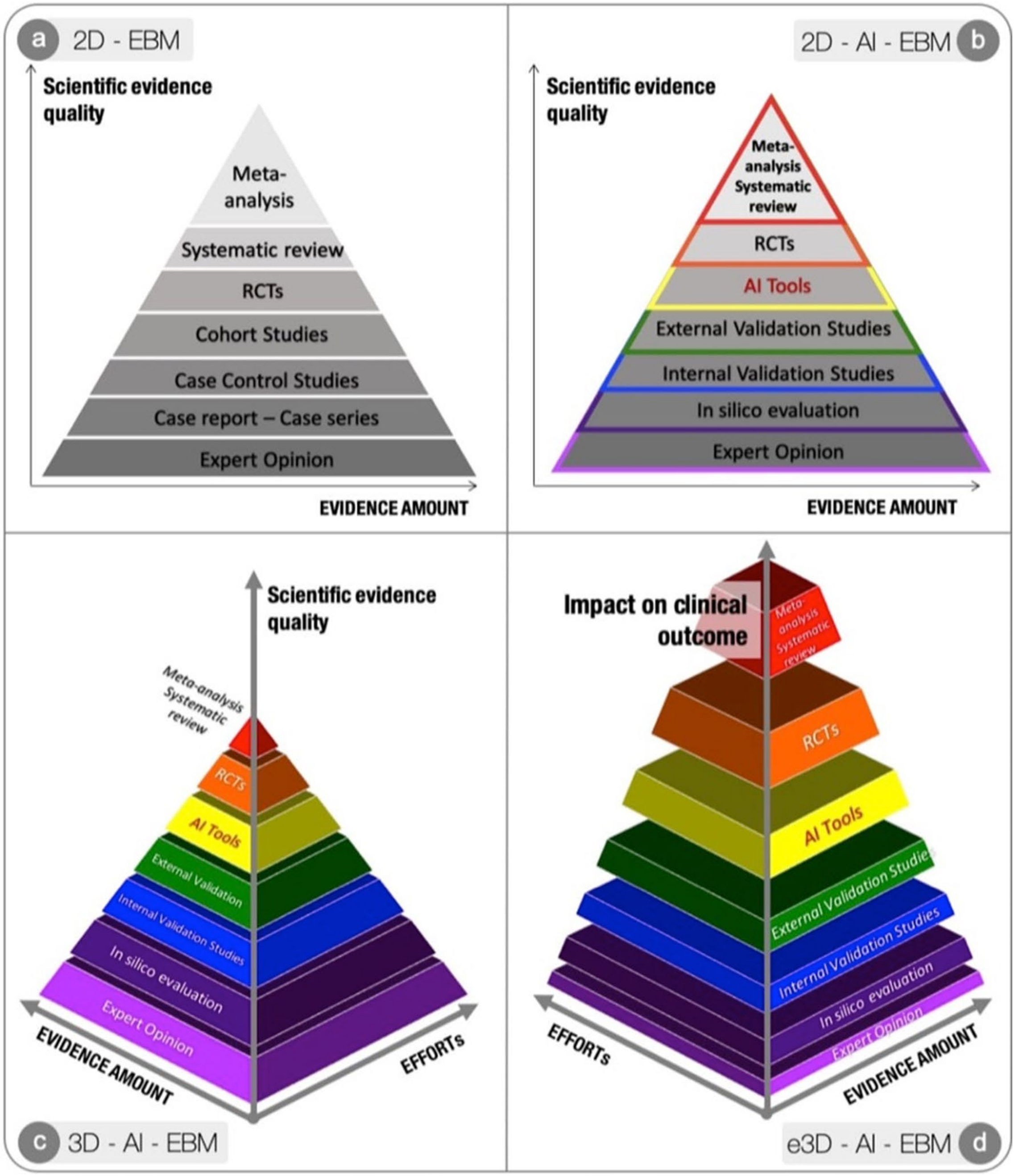

Fig. 1 Evidence pyramid metamorphosis in the era of disruption. **a** Traditional, 2D Pyramid of Evidence. Y-axis: scientific evidence quality; X-axis: evidence amount. **b** The Pyramid of Evidence following AI implementation, 2D-AI-EBM; Y-axis: scientific evidence quality; X-axis: evidence amount. **c** The 3D-AI-Pyramid of Evidence, obtained with a binocular vision that allows depth perception; Y-axis: scientific evidence quality; X-axis: evidence amount; Z-axis: efforts of implementation. **d** The enhanced, e3D-AI Pyramid of Evidence, that also analyze the impact on clinical outcome. Y-axis: impact on clinical outcome; X-axis: evidence amount; Z-axis: efforts of implementation (*AI* Artificial Intelligence)

developed to assess the risk of organ rejection in kidney transplantation to reduce its incidence [6] and thus to ensure a better quality of life for patients.

Of course, implementation and validation models are necessary in order to assess the impact and feasibility of the use of these technologies in daily clinical research and practice.

3 Reconsider bi-dimensional view: the depth of scientific evidence

Clinical practice, and so pyramid of evidence, should not be seen today as just a bi-dimensional figure (Fig. 1c). To have a 3D point-of-view, we need to have both our eyes working together: only in that way depth is perceived. On the one hand we have the evidence, on the other we have the complexity of conducting a more hierarchical study. This biocular vision thus gives rise to a three-dimensional pyramid. The peculiarity of 3D-view is to consider the exponentially efforts made to ascend only by a single step, related to the organizational and technical issues.

The management of such huge databases and the ethical-legal issues involved is one of the main concerns [12]. In fact, in many countries, aspects such as data protection or the management of multimedia clinical data have not yet been covered by new specific legislation or the necessary amendments have not been made to existing legislation.

Alongside these, there are some technological and logistic problems: in some hospital or research's structures are not present all new devices or interconnection systems between departments and servers. Where these are available, there must be adequate economic coverage to guarantee the management and maintenance of, for example, large databases for clinical data collected in real time. In other situations, they cannot be used for a lack of an adequate and performing internet line, leading to inequalities in access and use of these technologies. This results in upfront costs that are not supported by evidence of a real cost–benefit ratio, probably due to the lack of a consistent number of studies analysing the cost–benefit and effectiveness of AI implementation in a real context [13].

Where technological availability is not a constraint, using such devices in daily work life may still pose an objective challenge due to inadequate knowledge and unclear legislation, specifically in terms of responsibilities. For these reason different International Society like the American Medical Association in USA and Royal College of Physicians in Canada have recommended policies to integrate the use of AI and other new technologies into medical educational institutions [14].

Organizational challenges also arise due to the necessity of forming multidisciplinary teams consisting of data scientists, engineers, healthcare professionals, which can be challenging to implement, particularly in rural regions and countries with limited resources [12].

For all these elements, a two-dimensional pyramid, while accurate, does not fully capture the true essence of each stage. Technical and organizational limitations should always be taken into account at every stage of scientific research and the subsequent results into clinical practice.

4 Correlation between evidence quality and clinical outcome: an enhanced version of 3D-pyramid

The 3D-pyramid of evidence constructed in this manner requires the creation of an "enhanced" version in order to further augment its capabilities.

Each stage not only has a different volume and depth due to the efforts listed in the previous paragraph, but at each level the impact on clinical activities is amplified (Fig. 1d). As the pyramid ascends, the clinical effect increases gradually in direct relation to the size of each tier.

While the dissemination of artificial intelligence in the lower tiers of the pyramid has been easily accomplished, significant endeavor is required to generate studies such as RCTs or meta-analyses that can truly facilitate the dissemination of artificial intelligence in daily practice. It should not be surprising that expert opinions and theoretical models are much more numerous in literature than studies that aims to test AI tools with internal and external clinical validation.

There are several studies in the literature that, while reiterating the need for further research, express how the use of AI can be correlated with improved clinical practice. Abramoff et al. [15] describe an AI-based autonomous diagnostic system for the detection of diabetic retinopathy, which has passed Food and Drug Administration (FDA) scrutiny and has the potential to enhance primary care by incorporating specialist aspects. Other authors, such as Radenkovic et al. [3], on the other hand, point out a not insignificant aspect of clinical practice, namely the greater amount of time the health care professional can devote to patient care as a result of the reduction in bureaucracy required. Joshi et al. draw attention to how new technologies can reduce human error in the medical field, which is seen as both an improvement in collective health and a better use of available resources [16]. For example, they point out how mammogram reports integrated with biopsies can avoid up to 30 per cent of surgical procedures [17].

However, further medical research is needed to evaluate AI and other emerging technologies to improve diagnostic accuracy and clinical pathways for patients.

5 Conclusions and future directions

The pyramid remains incomplete without the integration of tools into clinical practice.

There are numerous key areas to focus on in order to achieve positive clinical outcomes.

Some of these include developing checklists to standardize the secure usage of AI [18], enacting ethical and privacy laws, working in multidisciplinary teams and providing medical school staff training [17]. In the meantime, efforts are to be made to produce high quality evidence supporting cost-benefits and cost-effectiveness of AI implementation in clinical practice, allowing healthcare professionals to rely on a solid evidence base when applying AI tools in real-world contexts.

Therefore, what do we do? Do we exclude these technologies? Do we discover how they can help us? After the first reasonable doubts, we must learn to live with them to achieve the maximum benefits for our patients. Only with this approach will it be possible to achieve truly personalized medicine based on each individual patient.

6 Core tip

- With the current technological revolution, evidence-based medicine is also changing.
- A progressive evolution of the evidence pyramid is proposed.
- Efforts are to be made in order to produce high quality evidence supporting cost-benefits and cost-effectiveness of AI implementation in clinical practice.

Author contributions VB, EO, FC and EB have equally made substantial contributions to all the following: (1) the conception and design of the study, acquisition of data, analysis, interpretation of data, (2) drafting the article and revising it critically, (3) final approval of the version to be submitted.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability Not applicable.

Code availability Not applicable.

Declarations

Competing interests Authors declare no competing interests.

References

1. Jones Emily. Digital disruption: artificial intelligence and international trade policy. Oxford Rev Econ Policy. 2023;39(1):70–84.
2. Pavaloia VD, Necula SC. Artificla intelligence as a disruptive technology- a systematic literature review. Eletronics. 2023;12(5):1102.
3. Radenkovic Dina, Keogh Sir Bruce, Maruthappu Mahiben. Data science in modern evidence-based medicine. J Royal Soc Med. 2019;112(12):493–4.

4. Maheshwari Kamal, Cywinski Jacek B, Papay Frank, Khanna Ashish K, Mathur Piyush. Artificial intelligence for perioperative medicine: perioperative intelligence. Anesth Analg. 2023. https://doi.org/10.1213/ANE.0000000000005952.
5. Zhong Yi, Lyu Yan-Ao-Hai, Song Yu, Gao Yu-jun, Mi Wei-Feng, Li Jian-Feng. The issue of evidence-based medicine and artificial intelligence. Asian J Psychiatry. 2023. https://doi.org/10.1016/j.ajp.2023.103627.
6. Burlacu Alexandru, Iftene Adrian, Busoiu Eugen, Cogean Dragos, Covic Adrian. Challenging the supremacy of evidence-based medicine through artificial intelligence: the time has come for a change of paradigms. Nephrol Dial Transplant. 2020;35:191–4.
7. Bellini Valentina, Carnà Emanuele Rafano, Russo Michele, Di Vincenzo Fabiola, Berghenti Matteo, Baciarello Marco, Bignami Elena. Artificial intelligence and anesthesia: a narrative review. Ann Transl Med. 2022;10(9):528.
8. Subbiah V. The next generation of evidence-based medicine. Nat Med. 2023;29(1):49–58.
9. Murad MH, Asi N, Alsawas M, Alahbad F. New evidence pyramid. Evid Based Med. 2016;21(4):125–7.
10. Bellini V, Coccolini F, Forfori F, Bignami EG. The artificial intelligence evidence-based medicine pyramid. World J Crit Care Med. 2023;12(2):89–91.
11. Williams Paul, BPharm, et al. In silico evaluation of a vancomycin dosing guideline among adults with serious infections. Ther Drug Monit. 2023. https://doi.org/10.1097/FTD.0000000000001102.
12. Cresswell Kathrin, Rigby Michael, Magrabi Farah, Scott Philip, Brender Jytte, Craven Catherine K, Wong Zoie Shui-Yee, Kukhareva Polina, Ammenwerth Elske, Georgiou Andrew, Meldock Stephanie, De Keizer Nicolette F, Nykanen Pirkko, Williams Robin. The need to strengthen the evaluation of the impact of artificial intelligence- based decision support systems on healthcare provision. Health Policy. 2023. https://doi.org/10.1016/j.healthpol.2023.104889.
13. Morris MX, Song EY, Rajesh A, Asaad M, Phillips BT. Ethical, legal and financial considerations of artificial intelligence in surgery. Am Surg. 2023;89(19):55–60.
14. Tolentino Raymond, Baradaran Ashkan, Gore Genevieve, Pluye Pierre, Abbasgholizadeh-Rahimi Samira. Curriculum frameworks and educational programs in artificial intelligence for medical students, residents, and practicing physicians: a scoping review protocol. JBI Evid Synth. 2023. https://doi.org/10.1112/JBIES-22-00374.
15. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. Digit Med. 2018;1:18.
16. Joshi Indra. Waiting for deep medicine. Lancet. 2019;393(10177):1193–4.
17. Otles E, James CA, Lomis KD, Woolliscroft JO. Teaching artificial intelligence as a fundamental toolset of medicine. Cell Rep Med. 2022;3(12):100824.
18. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, Arnaout R, Kohane IS, Saria S, Topol E, Obermeyer Z, Yu B, Butte AJ. Minimum information about clinical artifical intelligence modeling: the MI- CLAIM checklist. Nat Med. 2020;26(9):1320–4.