

Editorial

Giuseppe Lippi

Machine learning in laboratory diagnostics: valuable resources or a big hoax?

<https://doi.org/10.1515/dx-2019-0060>

Diagnostics shall still be considered a largely empirical science, where clinicians found their reasoning on a combination of medical history, physical examination and results of some informative diagnostic investigations, thus reflecting Sir William Osler's claim that "*medicine is a science of uncertainty and an art of probability*" [1]. This clear-cut definition implies that the diagnostic reasoning is strongly relying on human skills, which can be inherited, or can be garnered after many years of medical education and training [2]. In keeping with this preamble, human diagnostic reasoning is largely perfectible, so that innovative and effective tools that may help in improving human skills and more efficiently managing the high biological complexity characterizing the vast majority of human diseases would be enthusiastically welcomed.

Machine learning, a major branch of artificial intelligence (AI), is conventionally defined as a science where computer programs learn associations of predictive power from examples in data [3]. To put it simply, machine learning exploits predefined algorithms and statistical techniques, including those typically used in clinical medicine, for objective analysis of high-dimensional and multimodal biomedical data, with the ultimate scope – in laboratory diagnostics – of improving screening, identification, diagnosis, prognostication and therapeutic monitoring of human diseases [4].

A simple electronic search in Medline (interface PubMed) using the keyword "machine learning" produced as many as 25,323 hits at the time of publication of this editorial, with the time curve following an almost perfect exponential fitting during the past 30 years (Figure 1; $r = 0.989$; $p < 0.001$). By adding the keyword "laboratory medicine", the search output approximates 200 documents, following an even better exponential fitting during the past 10 years ($r = 0.994$; $p < 0.001$). This enormously boosted interest in machine learning, along with its potential applications in laboratory medicine, is generating great enthusiasm in the scientific community, as well as among laboratory professionals [5–7]. The possibility that machine learning may partially or entirely replace physician's brain is

indeed an intriguing perspective, whereby machines tend to be virtually foolproof when correctly trained. Unlike machines, the human brain is vulnerable to leaps, lapses and mistakes, especially when dealing with complex and multifaceted issues, such as integrating a large volume of demographic, clinical, environmental and instrumental information for making a final diagnosis. Although such a recent excitement in machine learning is substantially understandable, as the favorable support of AI to the clinical decision-making cannot be denied [5–7], many other reasons persuade me that its current value in laboratory diagnostics is maybe overrated.

The first, and probably most reasonable source of concern, is that machine learning is still dramatically dependent on the human counterpart. Because WE "humans" decide which is the pathologic condition that shall be targeted, WE "humans" identify which will be the study population and the study design (i.e. cross-sectional, retrospective, prospective or else), WE "humans" decide which are the demographical, clinical or diagnostic variables that shall be entered into the system, WE "humans" delineate which statistical tests shall be used, WE "humans" interpret results of the statistical analysis (e.g. area under the curve, sensitivity, specificity, negative predictive value, positive predictive value, Cox's regression analysis and so forth) and define whether these will be clinically valuable or not, WE "humans" allocate an error budget to predictive models and, last but not least, WE "humans" determine "if", "where", "when" and "how" models could be introduced into clinical practice. Irrespective of these shortcomings, it shall be also compellingly defined who is accountable for model validation before introduction into clinical practice, in analogy with the long and winding process that regulates translation of diagnostic tests from the bench to the bedside. This cannot be obviously left to software houses or individual scientists, but would need a deep engagement of scientific societies.

A second important aspect is that machines, even those perfectly trained, will be never capable of using Gestalt, intended as a partially innate and partially trainable perception, which allows to more or less consciously

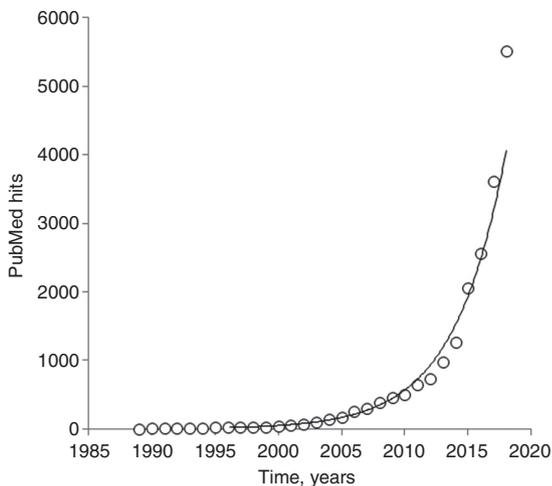


Figure 1: Exponential escalation of PubMed hits using the keyword “machine learning”.

integrate many diverse diagnostic elements [8]. Even the most sophisticated machine executes human commands, follows rigid schemes of (diagnostic) reasoning faster than humans, but will never be capable to extend analysis beyond the essential elements that humans have defined and incorporated. At least not in the foreseeable future.

Then, how can machine learning cope with personalized (laboratory) medicine? Precision (laboratory) medicine can be essentially defined as the process of making the right diagnosis, to the right patient, at the right time [9]. The progressive transition from conventional phenotypic testing to widespread use of innovative diagnostic strategies based on detection of genetic and/or epigenetic abnormalities hardly integrates with the low plasticity of machine learning. The clinical management of certain diseases – cancers are the most paradigmatic examples – is now strongly relying upon identification of highly specific (individual) somatic mutations which influence the clinical course, as well as the use of targeted and customized treatments. These disruptive technologies are now outlining the obsolescence of many diagnostic and therapeutic algorithms, thus making perhaps futile their application to machine learning [10].

The potential disruption of the vital laboratory-clinical liaison, which is at the very foundation of cost-effective usage of laboratory resources and patient-centered laboratory medicine, is another important drawback of machine learning. Most laboratory professionals are actively involved in counseling, consultancy and diagnostic stewardship, for actively promoting the most appropriate and effective use of laboratory resources [1]. How will this be possible after widespread implementation of

machine learning, when laboratory professionals will no longer be at the core of the diagnostic process?

A final remark, too often neglected, is who will be accountable for possible failures or mistakes of machine learning? Those who have developed the algorithm(s)? The software house? The healthcare organization? The single physician? Algorithms give the illusion of being foolproof and unbiased, but they have been developed by humans and trained on a specific set of data, which does not necessarily replicate local situations (e.g. populations are heterogeneous for many characteristics such as sex, age, ethnic origin, genetic background, co-morbidities, environment exposure and so forth). Needless to say, then, that the vast majority of human errors are cognitive in essence, whilst a machine error is typically repetitive, as it cannot be corrected without external (human) intervention.

In conclusion, I am rather challenged to establish whether machine learning will become a valuable future resource for laboratory medicine, or it will rather turn out as a big hoax. Neither hypothesis is perhaps completely correct or completely wrong. What is objectively clear is that the many current drawbacks of machine learning (Table 1) will not make it the panacea, nor AI will soon replace human reasoning throughout the diagnostic process. In the most worldwide celebrated manifest on AI and robotics, the famous novel writer and scientist Isaac Asimov (former Professor of Biochemistry at Boston University, by chance) has conceived the three law of robotics; the second of such laws clearly states that “*a robot must obey the orders given it by human beings except where such orders would conflict with the First Law of Robotics*” (Isaac Asimov. I, Robot. 1950). In this straightforward and almost unquestionable preamble, the use of the concept “must obey... by humans” is enlightening, whereby AI is (and will hopefully remain) subordinated to human intelligence.

Table 1: The current drawbacks of machine learning in laboratory medicine.

-
- Machines will always depend on (and obey to) human inputs
 - Who will be accountable for model(s) validation?
 - Machines need accurate data for generating trustable information
 - Machines do not consider unpredicted or unexpected variables
 - Machines will never use Gestalt perception
 - Models are not easily generalizable
 - Machines will hardly cope with personalized (laboratory) medicine
 - Who will be accountable for possible machine learning failures?
 - Machines will disrupt the vital laboratory-clinical liaison
-

Therefore, if a conclusion needs to be made at the end of this controversial editorial, it seems now unquestionable that machine learning is coming to fruition, as recently highlighted by Bergl et al. [11], that it will provide valuable support to the diagnostic reasoning, but its assistance shall be limited to advisory capacity, without unsupervised functioning and avoiding to completely surrogate human brain. If I was ill, I am not willing to be diagnosed (neither even cured) by robotics alone, at least for long time to come. I have little doubt that the vast majority of our patients are thinking alike.

Author contributions: The author has accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: None declared.

Employment or leadership: None declared.

Honorarium: None declared.

References

1. Plebani M, Laposata M, Lippi G. A manifesto for the future of laboratory medicine professionals. *Clin Chim Acta* 2019;489:49–52.
2. Lippi G, Cervellin G. From laboratory instrumentation to physician's brain calibration: the next frontier for improving diagnostic accuracy? *J Lab Precis Med* 2017;2:74.
3. Panch T, Szolovits P, Atun R. Artificial intelligence, machine learning and health systems. *J Glob Health* 2018;8:020303.
4. Sajda P. Machine learning for detection and diagnosis of disease. *Annu Rev Biomed Eng* 2006;8:537–65.
5. Gopal G, Suter-Crazzolara C, Toldo L, Eberhardt W. Digital transformation in healthcare – architectures of present and future information technologies. *Clin Chem Lab Med* 2019;57:328–35.
6. Cabitza F, Banfi G. Machine learning in laboratory medicine: waiting for the flood? *Clin Chem Lab Med* 2018;56:516–24.
7. Gruson D, Helleputte T, Rousseau P, Gruson D. Data science, artificial intelligence, and machine learning: opportunities for laboratory medicine and the value of positive regulation. *Clin Biochem* 2019;69:1–7.
8. Cervellin G, Borghi L, Lippi G. Do clinicians decide relying primarily on Bayesian principles or on Gestalt perception? Some pearls and pitfalls of Gestalt perception in medicine. *Intern Emerg Med* 2014;9:513–9.
9. Plebani M. Towards a new paradigm in laboratory medicine: the five rights. *Clin Chem Lab Med* 2016;54:1881–91.
10. Lippi G, Bassi A, Bovo C. The future of laboratory medicine in the era of precision medicine. *J Lab Precis Med* 2016;1:7.
11. Bergl PA, Wijesekera TP, Nassery N, Cosby KS. Controversies in diagnosis: contemporary debates in the diagnostic safety literature. *Diagnosis (Berl)* 2019. doi: 10.1515/dx-2019-0016 [Epub ahead of print].

Prof. Giuseppe Lippi, Section of Clinical Biochemistry, University Hospital of Verona, Piazzale LA Scuro, 37134 Verona, Italy, E-mail: giuseppe.lippi@univr.it