Opinion Paper

Anna Carobene*, Federico Cabitza, Sergio Bernardini, Raj Gopalan, Jochen K. Lennerz, Clare Weir and Janne Cadamuro

Where is laboratory medicine headed in the next decade? Partnership model for efficient integration and adoption of artificial intelligence into medical laboratories

https://doi.org/10.1515/cclm-2022-1030
Received October 12, 2022; accepted October 14, 2022; published online November 3, 2022

Abstract

Objectives: The field of artificial intelligence (AI) has grown in the past 10 years. Despite the crucial role of laboratory diagnostics in clinical decision-making, we found that the majority of AI studies focus on surgery, radiology, and oncology, and there is little attention given to AI integration into laboratory medicine.
Methods: We dedicated a session at the 3rd annual European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) strategic conference in 2022 to the topic of AI in the laboratory of the future. The speakers collaborated on generating a concise summary of the content that is presented in this paper.
Results: The five key messages are (1) Laboratory specialists and technicians will continue to improve the analytical portfolio, diagnostic quality and laboratory turnaround times; (2) The modularized nature of laboratory processes is amenable to AI solutions; (3) Laboratory sub-specialization continues and from test selection to interpretation, tasks increase in complexity; (4) Expertise in AI implementation and partnerships with industry will emerge as a professional competency and require novel educational strategies for broad implementation; and (5) regulatory frameworks and guidances have to be adopted to new computational paradigms.
Conclusions: In summary, the speakers opine that the ability to convert the value-proposition of AI in the laboratory will rely heavily on hands-on expertise and well designed quality improvement initiative from within laboratory for improved patient care.

Keywords: artificial intelligence; laboratory medicine; machine learning; performance metrics; robustness.

***Corresponding author: Anna Carobene**, Laboratory Medicine, IRCCS San Raffaele Scientific Institute, Via Olgettina 60, 20132 Milan, Italy, E-mail: carobene.anna@hsr.it
Federico Cabitza, IRCCS Ospedale Galeazzi - Sant'Ambrogio, Milan, Italy; and DISCo, Università Degli Studi di Milano-Bicocca, Milan, Italy
Sergio Bernardini, Unit of Laboratory Medicine, Tor Vergata University Hospital, Rome, Italy; and Department of Experimental Medicine, University of Tor Vergata, Rome, Italy
Raj Gopalan, Siemens Healthcare Diagnostics, Siemens Healthineers, Malvern, PA, USA
Jochen K. Lennerz, Department of Pathology, Center for Integrated Diagnostics, Harvard Medical School, Massachusetts General Hospital, Boston, MA, USA
Clare Weir, Sysmex Europe SE, Norderstedt, Germany
Janne Cadamuro, Department of Laboratory Medicine, Paracelsus Medical University Salzburg, Salzburg, Austria. https://orcid.org/0000-0002-6200-9831

Introduction

Healthcare without results from laboratory diagnostics has become unthinkable. Laboratory specialists and technicians have done an amazing job in the past decades, improving the analytical portfolio, quality and intra–laboratory turnaround time. Many laboratory processes are highly atomized, giving laboratory specialists the needed time resources to provide their expertise in test selection and interpretation to clinicians, a task, which currently is forced onto clinicians. However, the incredible improvement of laboratory processes due to automation and the speed with which the available tests have increased have led to a severe misuse (over- and underuse) of laboratory tests as well as incorrect interpretation of their results, including potential patient harm and financial impacts [1–4].

Personalized medicine demands dealing with individual patient samples, therefore systematic demand management solutions and clinical decision support systems (CDSS) are needed [5]. However, most laboratory

IT systems are incapable of modelling such solutions properly, which is why artificial intelligence (AI) algorithms have emerged during the past years [6].

Much has happened in the field of AI in the past 10 years, and there is no doubt that the trend in the number of scientific papers on medical AI studies follows an exponential growth [7]. However, searching by a medical speciality, we found out that the majority of studies focus on surgery, radiology, oncology, and although some few studies deal with AI integration into laboratory medicine.

In 2018, Cabitza et al. reported a lack of machine learning (ML)/AI studies using laboratory results. They expected that several such studies would be performed in the next years [8] as they partially verified with a subsequent literature review conducted only 3 years later [9]. The current trend of AI studies using medicine laboratory results is indeed growing exponentially and steadily so. Interestingly, there seems to be a gap between the number of studies listed in Scopus and PubMed databases. This gap can partly be explained by some IT journals or even conference papers appearing only in Scopus and not being indexed in the PubMed database. This observation is noteworthy because it suggests that laboratory data is used mainly by IT professionals rather than by laboratory specialists [10]. Confirming this assumption, recent reviews claim that even though the potential for ML models being applied in laboratory medicine is massive, only few tools in a wide number of medical studies have been translated into medical practice [11, 12].

Moreover, a recent opinion raises the problem to ensure appropriate rigour and transparency of ML-based methods studies, which will lead to their reproducibility, replicability, and clinical translation [13].

This article summarises the lectures held during a session at the 3rd European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) strategic conference, focussing on AI in laboratory medicine in real-life scenarios, including IT-and *In-vitro* Diagnostics (IVD) professionals.

How artificial intelligence could help the laboratory profession becoming more medical

Janne Cadamuro

The laboratory specialists focus on the analytical phase of the total testing process, aiming for higher quality, shorter turn around time (TATs), increased analytical accuracy, has led clinicians to think about the laboratory as a results producing facility. When categorizing all process steps within the pre-analytical, analytical and post-analytical phases into 'medical', 'organizational', and 'analytical', and combining these categories with the focus of most laboratories, it becomes evident that we have lost sight of the medical part of our profession (Figure 1). As mentioned above, contributing to patient care individually in unfeasible, which is why AI systems are needed aiding in test selection and interpretation. When fed some ground-truth data (data for the diagnostic work-up of the symptom or disease of interest, including all medical information, laboratory test results, and final diagnosis), these systems are able to calculate the most efficient and effective diagnostic pathway [14]. However, taking all variables into account, such as the patients pre-conditions, medications, physical examination, anamnesis, other diagnostic data and of course laboratory test results, is far more complex and error prone than image recognition models used in other medical disciplines. Using automated systems in medical care with limited accuracy levels may have devastating consequences. Jovičić et al. evaluated Mobile health applications, aiding patients in the interpretation of laboratory data and compared these with the opinions of laboratory professionals and confirmed the low utility of currently available laboratory medicine apps [15].

Therefore, in all the euphoria, we have to be cautious and careful and weigh the risks and benefits of such systems. Reports, such as the *Cambridge Analytica* scandal or the failed million-dollar project *IBM Watson*, are very good examples for how such systems may be misused or overrated [16, 17].

AI algorithms, aiding laboratory professionals in test selection and result interpretation will most probably not only make our profession more efficient, but more medical. Clinicians would be able to focus on their core expertise and laboratory professionals would answer clinical questions instead of reporting numbers. Therefore, the laboratory specialist should be open for new technology, rather than fearing the unknown.

Patient pathway and tailored diagnostics are becoming more prevalent in the clinical field

Clare Weir

(The opinions in this paragraph are expressly the view of the author and do not constitute the opinions or ideas of their employer).

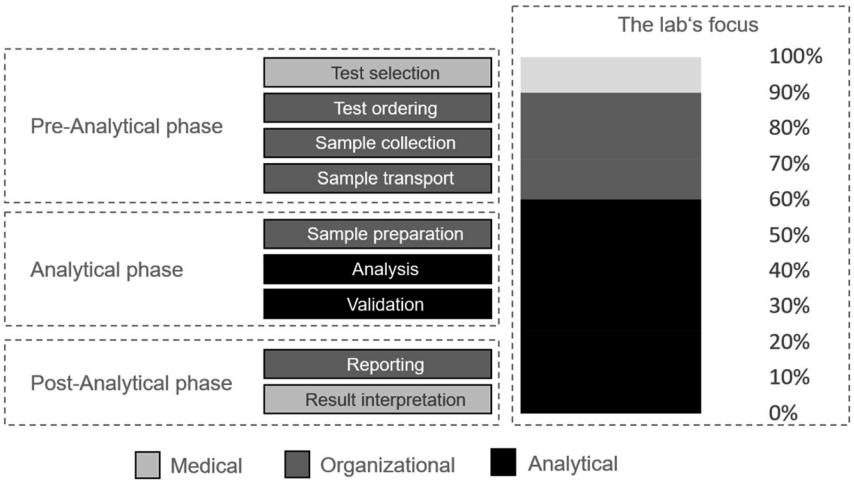

Figure 1: Categories of laboratory processes and the corresponding labs emphasis.

Patient pathway and tailored diagnostics are becoming more prevalent in the clinical field [18]. This ensures that the correct diagnosis and therapy is provided to the patient rather than a one for all system, which may not fit to the patient's history, symptoms and diagnosis. At this moment in time AI and ML is in its infancy in clinical laboratory practice and will probably not become mainstream for some years to come [19].

Working through the patient pathway from the first visit to a General Practitioner to monitoring of therapy in a haemato-oncology setting, there is only one area at the moment where ML and AI is applied and that is within the morphology section during the diagnostic process in both the haematology laboratory and flow cytometry laboratory (Figure 2). However, this does not mean that the morphologist is taken out of the picture; they are still required to validate the overall results.

In haematology, usually rule based algorithms are programmed into the middleware rather than AI models. These rules provide information on suspected abnormalities and provide suggestions of interpretive comments, possible reflex tests or alarms in case of critical values to the laboratory specialist. Although applying AI models to this area would be useful, the complexity and number of variables is currently too high to reach satisfying predictive values.

In flow cytometry, the use of samples such as bone marrow, cerebro-spinal fluid, synovial fluids and solid

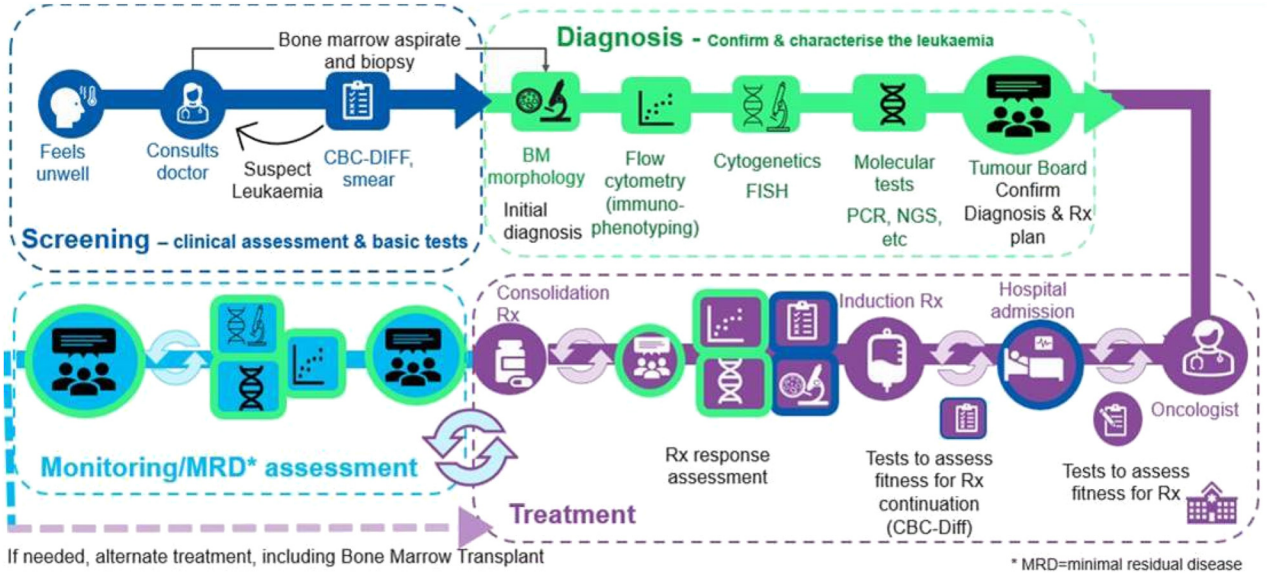

Figure 2: Illustrating the patient pathway from the perspective of a leukaemia or lymphoma[a]. [a]Picture provided by Dr. Marion Münster, Sysmex Europe SEU.

tumour tissue has to be acknowledged. Diagnostic errors by an AI/ML model may have devastating consequences, like insufficient, wrong or unnecessary treatment.

In an ideal setting, patient cases are discussed collaboratively among clinicians and laboratory specialists, each providing valuable expertise. Patient-specific suggestions, provided by AI/ML models, may improve this process in the future; however, currently this task is too complex to be done validly by such systems.

Finally, the patients' point of view needs to be considered from a psychological perspective. Medicine heavily relies on person-to-person interactions. Patients want to be treated as individuals, not as numbers or statistical data, which is why all professions in the medical field are one of the least likely to be atomized in the future [20].

The crucial role of laboratory diagnostics in clinical decision making in the era of artificial intelligence

Raj Gopalan

As mentioned above, although modern laboratories have become incredibly accurate and fast in sample processing and result delivery, test selection and result interpretation by clinicians continue to use a manual cognitive process, making them prone to human error and variation (Figure 3). The quality of this process depends on the clinician's training, experience, and diagnostic skills to determine the right differential diagnosis for the patient based on history, signs, symptoms, and physical examination. After a list of provisional differential diagnoses is formulated, clinicians must tap into their knowledge of diagnostics in laboratory medicine, as well as imaging, to choose the appropriate set of tests to help rule in and rule out the diagnoses on the list as well as to identify potential risks for future manifestation of specific diseases, such as cancer, liver or kidney diseases, or sepsis. When interpreting the results, clinicians must consider the contributions of the combined meaning of each of the laboratory test parameters and their clinical significance. Just scanning for abnormal result flags is not sufficient, since it does not assess important biochemical processes that may indicate underlying diseases. These tasks can be challenging, as the medical school educational curriculum does not typically include coursework on modern technology, such as AI and machine learning, to help future doctors manage these complex multidimensional problems [21, 22].

This results in many diagnostic errors and billions of dollars of malpractice claims each year [23, 24]. Therefore, the question remains: "Do clinical laboratories have a role in controlling skyrocketing healthcare costs in a post-COVID era? Can they help improve individual and population health by enabling early and accurate disease identification, clinical diagnosis, and treatment?" [24]. Even though laboratories' efficiency, accuracy, and consistency have improved over the years, delivering decision support to physicians, predominantly, has not expanded much beyond providing reference intervals and highlighting results that

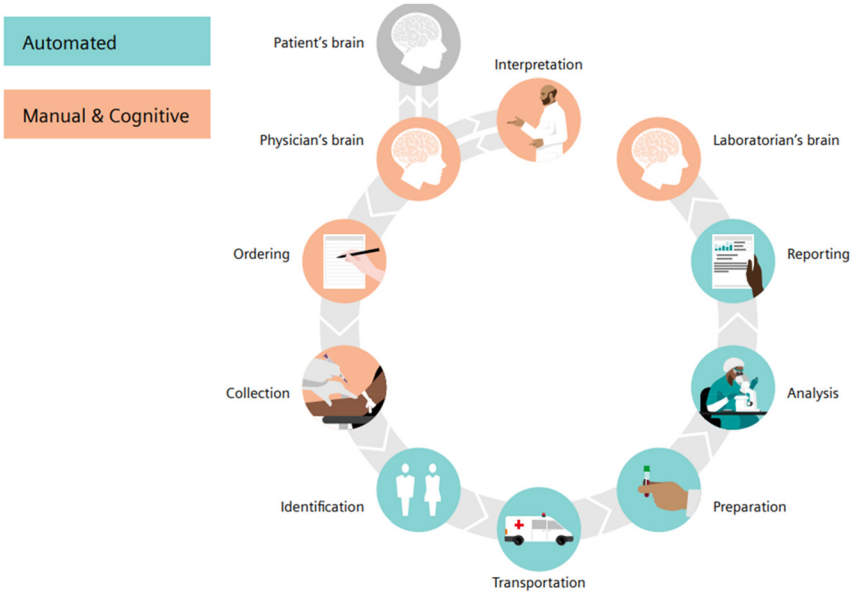

Figure 3: The brain-to-brain loop concept for laboratory testing[a]. [a]Adapted from Lundberg GD. Acting on significant laboratory results. JAMA. 1981;245:1762–3.

fall outside of those intervals [25]. Some laboratories are additionally providing delta-checks and reflex-testing algorithms.

Although these capabilities offer definitive value to clinicians, they are basic compared to those that laboratories can potentially offer. In recent times, and especially with the advent of molecular and genetic testing, the number of choices for laboratory testing has skyrocketed. The Choosing Wisely initiative started by the American Board of Internal Medicine (ABIM) to help clinicians order the appropriate tests is endorsed by more than 70 medical specialities. The goal of this initiative is to improve the safety and quality of patient care and reduce harm [26]. With the advent of electronic medical records (EMR), institutional laboratories have access to the initial, provisional diagnoses documented by clinicians. It seems logical that a future service provided by the laboratory would include recommendation of an optimal test battery that could be used to help to rule in the most probable diagnosis, rule out the unlikely ones, and identify the risks for life-threatening diseases. In addition, it may be possible to rank-order tests that are less expensive and covered by the patient's insurance. AI and ML technology can be used not only in the test-selection phase but also to help interpret multidimensional test results, providing likelihood scores that can be used to identify the risks for potential diseases and provide probability scores for subclinical diseases the patient may have, based on patients with a similar demographic and lab-test result profile from a large population of confirmed diagnosed cases [27]. CDSS can play a crucial role in tracking blood results and trends to monitor organ functions, thereby ensuring appropriate dosing to maintain patients within the most effective therapeutic spectrum. Thus, CDS creates the opportunity to pull data from multiple modalities, IVD, imaging, and the EMR; integrate that data; and present it visually to the multispecialty care team to enable comprehensive assessment of patients for optimal management decisions.

As we enter the post-COVID era, it is imperative to expand the traditional role of the clinical laboratory beyond its four walls to further benefit hospital units and clinics [28]. In addition to the decision-support information and services the laboratory can provide, there is also opportunity for more-direct engagement by laboratory specialists.

Laboratory directors who are trained physicians specializing in laboratory medicine have a unique role in working with other clinical specialists on care teams to evaluate, diagnose, and treat patients. With the advent of the EMR system and data interchange standards, laboratory directors can access complete medical data for patients and have the expertise to help clinicians order appropriate tests, interpret results, and guide further investigations.

"With each passing day, I place less value on accuracy"– some ideas to go beyond accuracy in evaluating machine-learning AI systems

Federico Cabitza

Laboratory medicine is no exception to the usage of computational classification systems created utilizing ML techniques in medicine for a variety of use cases and applications, such as diagnosis, prognosis, and risk stratification. However, though common and widely used in other domains, classification performance measures based on the concept of error (such as accuracy or the C-statistic) are unlikely to be meaningfully applicable to medicine.

Three primary reasons for distinct approaches for the evaluation of the efficacy of these decision support systems are here advocated. The first reason is the issue of replicability. Accuracy estimations are based on historical data with features that are often obtained from a single (or a few) institutions involved in the development of the ML model. Several studies have found that when used in different circumstances, even very accurate models report relevant drops in their accuracy [29]. The models must then be validated externally, using data from a diverse set of sources that are distinct (in terms of work habits and equipment) from those participating in the creation process [30, 31]. To achieve this goal, sound data similarity metrics must become more common and widely applied, allowing researchers to determine whether validation data are similar or different from training and test data, and thus whether accuracy scores are strongly correlated with similarity or not: this means focussing on robustness rather than accuracy. The second factor is the issue of noise [32]. Laboratory data are often thought of as good, reliable, and structured data; and rightly so, especially when compared to other types of health data, such as clinical or patient-reported data, which are influenced by observer variability to an extent that is often underestimated, if not completely ignored. ML processes, on the other hand, nearly entirely ignore the phenomena of analytical and biological variability [33]. This calls for data augmentation techniques and methods for creating synthetic data that are more indicative of the overall phenotypic complexity (and

variability) that defines the circumstances for which the system is asked to provide trustworthy advice or prediction. This is also a matter of robustness.

The third issue is that of meaning. Traditional error-based metrics distinguish between false-negative and false-positive errors, but they do not differentiate between instances on any dimension, such as relevance, diagnostic difficulty, rarity, or whether accurate results were also found to be interpretable by experts. Furthermore, these metrics do not weigh results in terms of prediction risk (that is, confidence score): as a result, estimates are produced by assuming not only that the cases are all equivalent for the sake of simplicity, but also that the decisions are all equal, even though they are based on very different probability estimates (and often with unknown or low calibration).

All of this calls for the creation of more reliable and comprehensive utility measurements that include predictive calibration. It is necessary to focus on interpretability, robustness, and utility rather than accuracy as a performance indicator in order to make AI support more reliable and trustworthy.

The importance of regulation of AI/ML in the laboratory

Jochen K. Lennerz

As discussed in the previous sections, AI has a huge potential to disrupt laboratory medicine soon. However, euphoria and promises cannot replace demonstration of clinical utility. It is noteworthy that clinical utility definitions vary across settings and use-cases [34]. When considering diagnostic testing – the promises of AI can be summarized as improvements in diagnostic testing quality. However, how can we capture the difficult topic of diagnostic test quality? Dr Lennerz proposed a nested conceptualization of diagnostic test quality that can be applied to the assessment and quantification of the quality impact when implementing AI models (Figure 4).

The conceptual starting point is the consideration that the AI model is integrated into the existing healthcare ecosystems. In the simplest case, the AI model (a computer program) becomes part of a specific diagnostic test (or a component of a test). The integration of the model as part of the diagnostic test is considered the first layer (=diagnostic test layer). Once accomplished, the model becomes an integral part of this specific laboratory testing process (also known as the laboratory procedures). The procedure that now entails an AI model is the second layer. With very few exceptions, most laboratories have multiple diagnostic testing procedures, and these may entail multiple models with their individual use cases, value propositions, and performance metrics (=value add). Each of these testing procedures is facing outward – externally towards the physician and the patient, which can be considered a diagnostic service. The diagnostic service offered by the laboratory (to a hospital or to outside partners) is considered the third layer.

In this conceptualization, the diagnostic quality is a composite of diagnostic test, diagnostic procedure, and diagnostic service. Consequently, the quality impact of AI models should be considered a function of the improvements related to the diagnostic test(s), the diagnostic procedure(s) and the diagnostic service(s) with their various intended use cases and value propositions.

How does this relate to regulatory issues? Now there are major challenges when encountering the portfolio of these hundreds and numerous elements that AI solution may touch. Among the major challenges and risks, issues regarding privacy, security, fairness, transparency and explainability, safety and performance, bias or third-party risks need to be considered [35, 36]. Regulations are complex and touch upon numerous aspects that culminate in compliance [37]. Documents and regulations in the field of AI regulation are manifold in and evolving [38, 39]. It is important to point out that regulations are made – and the field that challenges current regulation through the application of the scientific method (i.e., data) is called regulatory science [40]. Specifically, Regulatory science is the scientific discipline that evaluates and challenges current regulation; benefit versus risk assessment, and submission/approval strategies. Regulatory science uses a distinct terminology, aiming to describe what (intended use), who and why is doing something (indication of use), where and when it is done (context of use) and how is done (instructions of use). These definitions ultimately define the stringency during validation (=acceptance testing) and ongoing performance assessments (=proficiency testing) in clinical practice.

We have moved from diagnostics to advanced diagnostic technologies – we have now reached data science and AI. Where do we go next? We need to learn how to apply our high-quality data efficiently using computational technologies. However, when the complexity exceeds human capabilities, we need tools that help us streamline these divers functions and govern their risk and safety profiles. These tools include meaningful regulations.

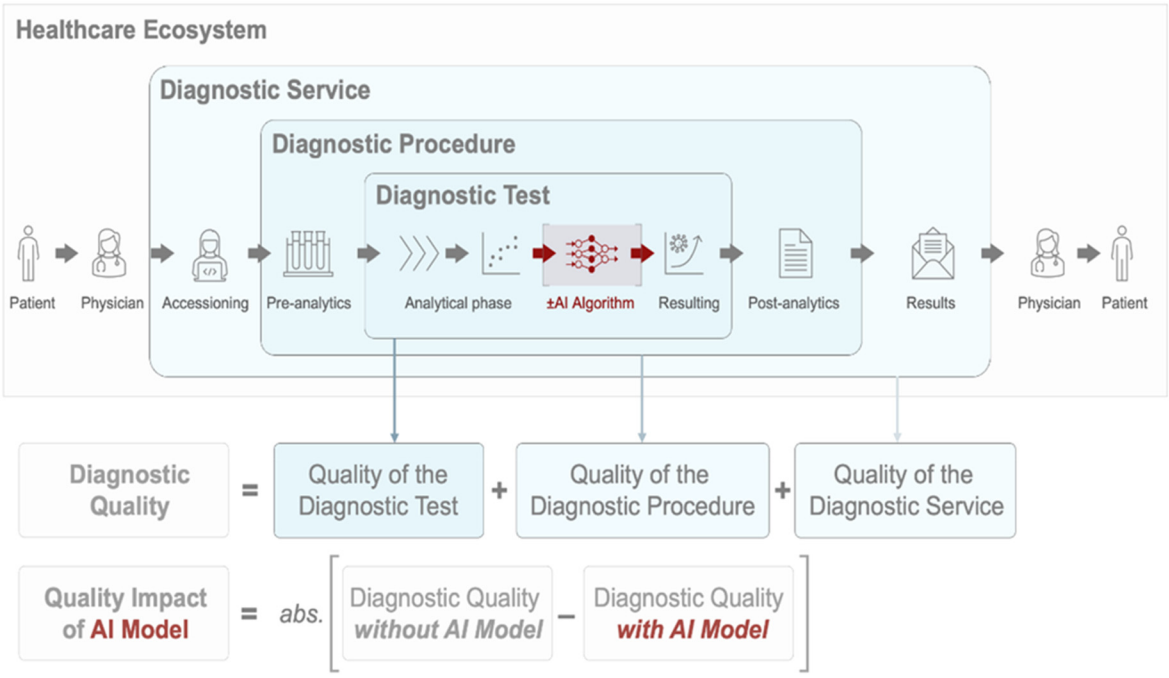

Figure 4: Conceptualization of a diagnostic AI model in the healthcare ecosystem. Each diagnostic test (first layer) is part of a specific set of operations collectively referred to as diagnostic procedure (second layer). Each laboratory typically has numerous diagnostic procedures. The diagnostic procedures interface with the external health care delivery system (third layer, diagnostic service layer). The diagnostic quality can be seen as the combination of the diagnostic quality of the diagnostic test, procedure, and service. Considering the deployment of AI models, the diagnostic quality impact of an AI model can be expressed as the absolute (abs.) difference between the quality with or without the AI model. For simplicity of the conceptualization, the AI model is exemplarily depicted in the diagnostic test layer; however, AI models can be implemented in other and/or multiple layers.

In other words, what is next in laboratory medicine? We should consider the ability to actively influence laboratory regulation as an emerging field in laboratory medicine. Rather than receiving the rules, we can actively contribute, share our knowledge, and help shape our field.

Final remarks

As was widely stated by all speakers in this session, artificial intelligence is a disruptive technology that must be adapted to, and adopted by the laboratory specialists profession, given its increasing availability and application in laboratory medicine [41].

A recent survey revealed a lack of specific knowledge on the subject of Big Data and Artificial Intelligence among Italian clinical laboratory professionals [42]. Additionally, it raised concerns regarding infrastructure prerequisites, a general lack of hardware and software infrastructures, dearth of personal PCs, lack of corporate Wi-Fi networks, and a low level of subjective satisfaction with regard to both software and hardware equipment.

In another survey on the value of AI in laboratory medicine, recently conducted in the United States among laboratory stakeholders [19], the perceived value of AI observed does not differ from that of the general population; similar results concerning infrastructure prerequisites, were reported as deficient by both questionnaires. Additionally, in the aforementioned studies, the majority of respondents said they were unsure about using AI in their businesses or that they would never do so. The question remains whether this behaviour based on the fear of replacement by this disrupting technology, or if there are other underlying causes.

In case of the former reason, this misgiving should be overcome by the awareness that AI will become an inevitable part of laboratory testing and it should be considered as improvement in diagnostic quality [6]. Effective collaboration between data scientists and clinicians is essential in obtaining clinically useful models. Additionally, co-operation with the IVD industry is an important prerequisite in obtaining useful ML models, as in addition to digital image analysis (hematopatology, microbiology, urine sediment analysis etc.), recently prognostic and

diagnostic algorithms are being developed, using numerical laboratory data.

In all the euphoria about this emerging technology, we should not forget about its limitations, avoiding jeopardizing patient safety. Therefore, laboratory specialists are now tasked to develop the tools, standards, and experimental approaches, improving assessment of safety, efficacy, quality and performance of ML/AI models used in patient care.

Nevertheless, clinicians should spend time learning the fundamentals of these new technologies in order to evaluate clinical trial opportunities. Indeed, the involvement of laboratory specialists is crucial to ensure that laboratory data are sufficiently available and conscientiously incorporated into strong, safe and clinically successful scientific projects, aiming to improve patient safety through higher quality and laboratory diagnostics [9].

Research funding: None declared.
Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: Authors state no conflict of interest.
Informed consent: Not applicable.
Ethical approval: Not applicable.

References

1. Cadamuro J, Gaksch M, Wiedemann H, Lippi G, von Meyer A, Pertersmann A, et al. Are laboratory tests always needed? Frequency and causes of laboratory overuse in a hospital setting. Clin Biochem 2018;54:85–91.
2. Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaout R. The landscape of inappropriate laboratory testing: a 15-year meta-analysis. PLoS One 2013;8:e78962.
3. Cadamuro J, Simundic AM, von Meyer A, Haschke-Becher E, Keppel MH, Oberkofler H, et al. Diagnostic workup of microcytic anemia. Arch Pathol Lab Med 2022. https://doi.org/10.5858/arpa.2021-0283-OA [Epub ahead of print].
4. Mrazek C, Lippi G, Keppel MH, Felder TK, Oberkofler H, Haschke-Becher E, et al. Errors within the total laboratory testing process, from test selection to medical decision-making – a review of causes, consequences, surveillance and solutions. Biochem Med (Zagreb) 2020;30:020502.
5. Hoffmann GE, Aufenanger J, Födinger M, Cadamuro J, von Eckardstein A, Kaeslin-Meyer M, et al. Benefits and limitations of laboratory diagnostic pathways. Diagnosis 2014;1:269–76.
6. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25:44–56.
7. Mesko B, Gorog M. A short guide for medical professionals in the era of artificial intelligence. NPJ Digit Med 2020;3:126.
8. Cabitza F, Banfi G. Machine learning in laboratory medicine: waiting for the flood? Clin Chem Lab Med 2018;56:516–24.
9. Ronzio L, Cabitza F, Barbaro A, Banfi G. Has the flood entered the basement? A systematic literature review about machine learning in laboratory medicine. Diagnostics (Basel) 2021;11:372.
10. Carobene A, Milella F, Famiglini L, Cabitza F. How is test laboratory data used and characterised by machine learning models? A systematic review of diagnostic and prognostic models developed for COVID-19 patients using only laboratory data. Clin Chem Lab Med 2022. https://doi.org/10.1515/cclm-2022-0182 [Epub ahead of print].
11. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med 2022;28:31–8.
12. Herman DS, Rhoads DD, Schulz WL, Durant TJS. Artificial intelligence and mapping a new direction in laboratory medicine: a review. Clin Chem 2021;67:1466–82.
13. Haymond S, Master SR. How can we ensure reproducibility and clinical translation of machine learning applications in laboratory medicine? Clin Chem 2022;68:392–5.
14. Cadamuro J. Rise of the machines: the inevitable evolution of medicine and medical laboratories intertwining with artificial intelligence-A narrative review. Diagnostics (Basel) 2021;11:1399–416.
15. Jovicic S, Siodmiak J, Alcorta MD, Kittel M, Oosterhuis W, Aakre KM, et al. Quality benchmarking of smartphone laboratory medicine applications: comparison of laboratory medicine specialists' and non-laboratory medicine professionals' evaluation. Clin Chem Lab Med 2021;59:693–9.
16. Strickland E. How IBM Watson overpromised and underdelivered on AI health care. https://spectrum.ieee.org/biomedical/diagnostics/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care [Accessed 30 June 2021].
17. Wong JC. The Guardian – the Cambridge analytica scandal changed the world – but it didn't change Facebook. https://www.theguardian.com/technology/2019/mar/17/the-cambridge-analytica-scandal-changed-the-world-but-it-didnt-change-facebook [Accessed 30 June 2021].
18. Chalmers D, Nicol D, Otlowski M, Critchley C. Personalised medicine in the genome era. J Law Med 2013;20:577–94.
19. Paranjape K, Schinkel M, Hammer RD, Schouten B, Nannan Panday RS, Elbers PWG, et al. The value of artificial intelligence in laboratory medicine. Am J Clin Pathol 2021;155:823–31.
20. Whitehouse M, Rojanasakul M. Find out if your job will Be automated. https://www.bloomberg.com/graphics/2017-job-risk/ [Accessed 27 July 2022].
21. Baron JM, Dighe AS, Arnaout R, Balis UJ, Black-Schaffer WS, Carter AB, et al. The 2013 symposium on pathology data integration and clinical decision support and the current state of field. J Pathol Inf 2014;5:2.
22. Smith BR, Kamoun M, Hickner J. Laboratory medicine education at U.S. medical schools: a 2014 status report. Acad Med 2016;91:107–12.
23. Newman-Toker DE, editor. Diagnostic value: the economics of high-quality diagnosis and value-based perspective on diagnostic innovation. Modern healthcare annual patient safety and quality virtual conference; 2015.
24. Johns Hopkins Medicine. Diagnostic errors more common, costly and harmful than treatment mistakes. https://www.hopkinsmedicine.org/news/media/releases/diagnostic_errors_more_common_costly_and_harmful_than_treatment_mistakes [Accessed 27 July 2022].

25. Pillay TS. Artificial intelligence in pathology and laboratory medicine. J Clin Pathol 2021;74:407–8.
26. Baird GS. The choosing Wisely initiative and laboratory test stewardship. Diagnosis (Berl) 2019;6:15–23.
27. Guncar G, Kukar M, Notar M, Brvar M, Cernelc P, Notar M, et al. An application of machine learning to haematological diagnosis. Sci Rep 2018;8:411.
28. Singh V, Kamaleswaran R, Chalfin D, Buno-Soto A, San Roman J, Rojas-Kenney E, et al. A deep learning approach for predicting severity of COVID-19 patients using a parsimonious set of laboratory markers. iScience 2021;24:103523.
29. Cabitza F, Campagner A, Ferrari D, Di Resta C, Ceriotti D, Sabetta E, et al. Development, evaluation, and validation of machine learning models for COVID-19 detection based on routine blood tests. Clin Chem Lab Med 2020;59:421–31.
30. Cabitza F, Campagner A, Soares F, Garcia de Guadiana-Romualdo L, Challa F, Sulejmani A, et al. The importance of being external. methodological insights for the external validation of machine learning models in medicine. Comput Methods Progr Biomed 2021;208:106288.
31. Campagner A, Carobene A, Cabitza F. External validation of machine learning models for COVID-19 detection based on complete blood count. Health Inf Sci Syst 2021;9:37.
32. Cabitza F, Campagner A, Albano D, Aliprandi A, Bruno A, Chianca V, et al. The elephant in the machine: proposing a new metric of data reliability and its application to a medical case to assess classification reliability. Appl Sci-Basel 2020;10: 4014–32.
33. Campagner A, Famiglini L, Carobene A, Cabitza F. Everything is varied: the surprising impact of individual variation on ML robustness in medicine. arXiv preprint arXiv:221004555; 2022.
34. Pritchard D, Goodman C, Nadauld LD. Clinical utility of genomic testing in cancer care. JCO Precis Oncol 2022;6:e2100349.
35. Buehler M, Dooley R, Grennan L, Singla A. A systematic approach to identifying and prioritizing AI risks can help organizations effectively target mitigation efforts. https://www.mckinsey.com/business-functions/quantumblack/our-insights/getting-to-know-and-manage-your-biggest-ai-risks [Accessed 27 July 2022].
36. Pennestri F, Banfi G. Artificial intelligence in laboratory medicine: fundamental ethical issues and normative key-points. Clin Chem Lab Med 2022. https://doi.org/10.1515/cclm-2022-0096 [Epub ahead of print].
37. Huang R, Lasiter L, Bard A, Quinn B, Young C, Salgado R, et al. National maintenance cost for precision diagnostics under the verifying accurate leading-edge in vitro clinical test development (VALID) act of 2020. JCO Oncol Pract 2021;17:e1763–73.
38. Marble HD, Huang R, Dudgeon SN, Lowe A, Herrmann MD, Blakely S, et al. A regulatory science initiative to harmonize and standardize digital pathology and machine learning processes to speed up clinical innovation to patients. J Pathol Inf 2020;11:22.
39. International Organization for Standardization. ISO/IEC 23053: 2022 – framework for artificial intelligence (AI) systems using machine learning (ML); 2022.
40. Lennerz JK, Marble HD, Lasiter L, Poste G, Sirintrapun SJ, Salgado R. Do not sell regulatory science short. Nat Med 2021;27: 573–4.
41. Cadamuro J. Disruption vs. evolution in laboratory medicine. Current challenges and possible strategies, making laboratories and the laboratory specialist profession fit for the future. Clin Chem Lab Med 2022. https://doi.org/10.1515/cclm-2022-0620 [Epub ahead of print].
42. Bellini C, Padoan A, Carobene A, Guerranti R. A survey on Artificial Intelligence and Big Data utilisation in Italian clinical laboratories. Clin Chem Lab Med 2022. https://doi.org/10.1515/cclm-2022-0680 [Epub ahead of print].