Clinical Microbiology and Infection 26 (2020) 1289–1290

Contents lists available at ScienceDirect

Clinical Microbiology and Infection

journal homepage: www.clinicalmicrobiologyandinfection.com

Editorial

Digitalization, clinical microbiology and infectious diseases

In recent years, digitalization and artificial intelligence have made tremendous progress. In medicine, data-driven technologies are especially applicable in areas with a high degree of automation and standardization of data [1,2]. Substantial advances have as well been reported in clinical microbiology, but their translation into routine application remains a long process with several technical and regulatory hurdles. Some of the low-hanging fruits for diagnostics scenarios include (i) dashboards to interconnect and visualize microbiology data [3,4], (ii) automated analysis of images such as microscopy slides [5] or agar plates [6,7] and (iii) association of genome sequences and proteomic profiles with pathogen phenotypes [8,9]. Clinical applications require standardized data formats, ontologies with an interoperable information technology environment [10], infrastructure with sufficient storage and computational capacity, and technical expertise to address the needs of microbiologists and infectious diseases experts.

In the present themed issue, Luz et al. summarize machine learning algorithms for the analysis of routine electronic health records. The authors identified 52 studies covering various aspects of infectious disease management including sepsis, hospital-acquired and surgical site infections, and microbiological test results. The heterogeneity of machine learning algorithms ranged from logistic regression, random forest, support vector machines to artificial neural networks. A key gap is the lack of essential information on data handling [11]. Pfeiffer-Smadja et al. ask if the time has come for machine learning in routine practice of clinical microbiology [12]. In 97 studies, the data sources used were highly diverse ranging from genomic data and microscopic images to mass spectrometry. Almost 40% of studies were from low- and middle-income countries—highlighting the opportunities that digitalization and digital biomarkers have to offer considering decreasing costs and cloud-based services [13]. However, digital biomarkers also require validation in clinical studies to show their impact on relevant outcomes. Lacking standardized data and algorithms poses an important challenge for reproduction and validation studies [14]. As a result of issues in data handling, two prominent published coronavirus disease 2019 (COVID-19) articles were recently retracted [15,16]. Journals clearly need standards for data and code sharing. The FAIR principles provide an excellent guidance [17]. Although software code and tools are often shared on GitHub (github.com) [18], the details provided are often limited with missing explanatory code books or instructions. Proper data and code-handling policies should be part of the new research quality standard and will allow independent validation of machine learning algorithms and data sets.

Smith and Kirby report on applications in modern image analysis [19]. Machine-learning-based image analysis may revolutionize microscopy for classical Gram stains, ova and parasite preparation, and histopathological slides. For example, a neural network could categorize Gram stains from positive blood cultures with remarkable precision into Gram positives/negatives and cocci/rods [5]. Of note, state-of-the-art infrastructure to generate high-quality images, data storage and processing may be required. However, smartphone devices can bridge the technology gaps [20,21]. Similarly, based on pattern recognition, single bacterial colonies growing on agar plates can be categorized or even identified [6,7]. Both applications, automated microscopy and agar plate inspection, are likely to radically change the workflow in modern diagnostic laboratories [22]. Perhaps parallel to how we have embraced matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass-spectrometry for identification, making biochemical tests almost superfluous [23]. However, there may be potential for extracting additional information from MALDI-TOF spectra. Weis and colleagues look into this key technology [24] and summarize algorithms to link spectral profiles to microbiological phenotypes. In their review, 36 studies using machine learning for species identification and antibiotic susceptibility testing were identified. Most commonly used machine learning techniques included support vector machines, genetic algorithms, artificial neural networks and quick classifiers. Within the studies identified, a wide range of qualities were noted and only four studies validated their findings [24].

All authors highlight the need for validated algorithms. Validation is also a key point in the regulatory process and impacts reimbursement. From May 2021, the medical device and *in vitro* medical device regulations of Europe will steer software with a diagnostic, monitoring or therapeutic purpose (http://ec.europa.eu/growth/sectors/medical-devices/regulatory-framework/), forming the basis for CE labelling including machine-learning-based algorithms in clinical microbiology. Both academia and industry will benefit from standards in data and code handling as this process will support validation and further build trust in computational models and methods [25,26]. A process additionally fuelled by (i) well-designed clinical studies and (ii) cross-validation to known and well-established statistical approaches. Ethical and legal aspects should also be raised if such algorithms are to be integrated into personalized and public health medicine [27]. As illustrated, during the COVID-19 crisis, multiple models have predicted different outcomes [28,29] of, for example, fatality rates and impact of the lockdown. In public health emergencies high-quality real-time data must be available in machine-readable formats for the scientific community. Such infrastructure for public health monitoring needs to be further developed. If public health decisions rely on such models, in return models should to be validated in a similar way

https://doi.org/10.1016/j.cmi.2020.06.031

to algorithms in personalized medicine because the impact for the general population and economics is significant.

Clearly, an interesting and challenging time for clinical microbiology and infectious disease is ahead. Standards in data and code handling are a first step, which will allow us to use the opportunities of digitalization and machine learning to improve diagnostics and patient care.

References

[1] Bailey AL, Ledeboer N, Burnham CD. Clinical microbiology is growing up: the total laboratory automation revolution. Clin Chem 2019;65:634–43.

[2] Ezewudo M, et al. Integrating standardized whole genome sequence analysis with a global *Mycobacterium tuberculosis* antibiotic resistance knowledgebase. Sci Rep 2018;8:15382.

[3] Graber CJ, et al. Decreases in antimicrobial use associated with multihospital implementation of electronic antimicrobial stewardship tools. Clin Infect Dis 2019. https://doi.org/10.1093/cid/ciz941. epub ahead of print.

[4] Hebert C, Flaherty J, Smyer J, Ding J, Mangino JE. Development and validation of an automated ventilator-associated event electronic surveillance system: a report of a successful implementation. Am J Infect Contr 2018;46:316–21.

[5] Smith KP, Kang AD, Kirby JE. Automated interpretation of blood culture Gram stains by use of a deep convolutional neural network. J Clin Microbiol 2018;56:e01521. 17.

[6] Croxatto A, et al. Towards automated detection, semi-quantification and identification of microbial growth in clinical bacteriology: a proof of concept. Biomed J 2017;40:317–28.

[7] Van TT, Mata K, Dien Bard J. Automated detection of *Streptococcus pyogenes* pharyngitis by use of colorex strep A CHROMagar and WASPLab artificial intelligence chromogenic detection module software. J Clin Microbiol 2019;57:e00811–9.

[8] Jamal S, et al. Artificial intelligence and machine learning based prediction of resistant and susceptible mutations in *Mycobacterium tuberculosis*. Sci Rep 2020;10:5487.

[9] Lupolova N, Lycett SJ, Gally DL. A guide to machine learning for bacterial host attribution using genome sequence data. Microb Genom 2019;5. https://doi.org/10.1099/mgen.0.000317.

[10] Gansel X, Mary M, van Belkum A. Semantic data interoperability, digital medicine, and e-health in infectious disease management: a review. Eur J Clin Microbiol Infect Dis 2019;38:1023–34.

[11] Luz CF, et al. Machine learning in infection management using routine electronic health records: tools, techniques, and reporting of future technologies. Clin Microbiol Infect 2020;26:1291–9. https://doi.org/10.1016/j.cmi.2020.02.003.

[12] Pfeiffer-Smadja N, et al. Machine learning in the clinical microbiology laboratory: has the time come for routine practice? Clin Microbiol Infect 2020;26:1300–9. https://doi.org/10.1016/j.cmi.2020.02.006.

[13] Karim MR, et al. Improving data workflow systems with cloud services and use of open data for bioinformatics research. Brief Bioinform 2018;19:1035–50.

[14] Vandenberg O, et al. Consolidation of clinical microbiology laboratories and introduction of transformative technologies. Clin Microbiol Rev 2020;33. https://doi.org/10.1128/CMR.00057-19.

[15] Mehra MR, Desai SS, Kuy S, Henry TD, Patel AN. Retraction: cardiovascular disease, drug therapy, and mortality in Covid-19. N Engl J Med 2020. https://doi.org/10.1056/NEJMoa2007621. N Engl J Med, doi:10.1056/NEJMc2021225.

[16] Mehra MR, Desai SS, Ruschitzka F, Patel AN. RETRACTED: hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. Lancet 2020. https://doi.org/10.1016/S0140-6736(20)31180-6.

[17] Corpas M, Kovalevskaya NV, McMurray A, Nielsen FGG. A FAIR guide for data providers to maximise sharing of human genomic data. PLoS Comput Biol 2018;14:e1005873.

[18] Hendriksen RS, et al. Using genomics to track global antimicrobial resistance. Front Public Health 2019;7:242.

[19] Smith KP, Kirby JE. Image analysis and artificial intelligence in infectious disease diagnostics. Clin Microbiol Infect 2020;26:1318–23. https://doi.org/10.1016/j.cmi.2020.03.012.

[20] Linares M, et al. Collaborative intelligence and gamification for on-line malaria species differentiation. Malar J 2019;18:21.

[21] Perkel JM. Pocket laboratories. Nature 2017;545:119–21.

[22] Cherkaoui A, et al. Implementation of the WASPLab and first year achievements within a university hospital. Eur J Clin Microbiol Infect Dis 2020. https://doi.org/10.1007/s10096-020-03872-1.

[23] Angeletti S, Ciccozzi M. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry in clinical microbiology: an updating review. Infect Genet Evol 2019;76:104063.

[24] Weis CV, Jutzeler CR, Borgwardt K. Machine learning for microbial identification and antimicrobial susceptibility testing on MALDI-TOF mass spectra: a systematic review. Clin Microbiol Infect 2020;26:1310–7. https://doi.org/10.1016/j.cmi.2020.03.014.

[25] Cabitza F, Campagner A, Balsano C. Bridging the "last mile" gap between AI implementation and operation: "data awareness" that matters. Ann Transl Med 2020;8:501.

[26] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25:44–56.

[27] Watson DS, et al. Clinical applications of machine learning algorithms: beyond the black box. BMJ 2019;364:l886.

[28] Davies NG, et al. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. Lancet Public Health 2020. https://doi.org/10.1016/S2468-2667(20)30133-X.

[29] Flaxman S, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature 2020. https://doi.org/10.1038/s41586-020-2405-7.

A. Egli*

Clinical Bacteriology and Mycology, University Hospital Basel, Basel, Switzerland

Applied Microbiology Research, Department of Biomedicine, University of Basel, Basel, Switzerland

* Corresponding author: A. Egli, Clinical Bacteriology and Mycology, University Hospital Basel, Petersgraben 4, 4031, Basel, Switzerland.
E-mail address: adrian.egli@usb.ch.

15 June 2020
Available online 2 July 2020

Editor: L. Leibovici