



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Review

Digital microbiology

A. Egli^{1,2,*}, J. Schrenzel³, G. Greub⁴¹ Clinical Bacteriology and Mycology, University Hospital Basel, Basel, Switzerland² Applied Microbiology Research, Department of Biomedicine, University of Basel, Basel, Switzerland³ Laboratory of Bacteriology, University Hospitals of Geneva, Geneva, Switzerland⁴ Institute of Medical Microbiology, University Hospital Lausanne, Lausanne, Switzerland

ARTICLE INFO

Article history:

Available online 27 June 2020

Editor: L Leibovici

Keywords:

Analytics
Artificial intelligence
Diagnostics
Digitalization
Image analysis
Interoperability
Microbiology
Pre-analytics
Post-analytics
Quality

ABSTRACT

Background: Digitalization and artificial intelligence have an important impact on the way microbiology laboratories will work in the near future. Opportunities and challenges lie ahead to digitalize the microbiological workflows. Making efficient use of big data, machine learning, and artificial intelligence in clinical microbiology requires a profound understanding of data handling aspects.

Objective: This review article summarizes the most important concepts of digital microbiology. The article gives microbiologists, clinicians and data scientists a viewpoint and practical examples along the diagnostic process.

Sources: We used peer-reviewed literature identified by a PubMed search for digitalization, machine learning, artificial intelligence and microbiology.

Content: We describe the opportunities and challenges of digitalization in microbiological diagnostic processes with various examples. We also provide in this context key aspects of data structure and interoperability, as well as legal aspects. Finally, we outline the way for applications in a modern microbiology laboratory.

Implications: We predict that digitalization and the usage of machine learning will have a profound impact on the daily routine of laboratory staff. Along the analytical process, the most important steps should be identified, where digital technologies can be applied and provide a benefit. The education of all staff involved should be adapted to prepare for the advances in digital microbiology. **A. Egli, Clin Microbiol Infect 2020;26:1324**

© 2020 The Authors. Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Without doubt, digital technologies will shape our lives in the upcoming years: from personal assistants [1], internet-connected devices and bodies [2,3] including smart phone technologies [4], self-driving cars and drones [5,6], to algorithms for self-improvement [7,8]. Digitalization and artificial intelligence (see [Supplementary Table S1](#) for a glossary) generate high expectations in healthcare [9]. These expectations are fuelled by an increasing demand to optimize quality and lower costs. According to the Organisation for Economic Co-operation and Development (OECD), health spending in 2017 as a share of the gross domestic product (GDP) was 8.8% on average, corresponding to USD 3857 per capita per year. Many countries have observed a substantial increase, with

healthcare costs more than doubling over the past 10 years (<https://data.oecd.org>). Therefore, various stakeholders have great hope for the magic bullet of digitalization to control or even lower healthcare-associated costs. General aspects of digitalization in medicine have been recently reviewed elsewhere [9–11]. Nevertheless, digitalization will lead to a significant optimization of healthcare-associated processes and improvement of quality needs to be seen. Accordingly, there will be a higher demand for high-quality digital laboratory and specifically also microbiological data in diagnostics [12] in order to (a) use machine learning algorithms for optimization of the treatment indication and prediction of prognosis, and (b) as information sources to monitor and document the quality and impact of medical interventions.

The increasing need for microbiological digital data is also an opportunity for microbiologists and other laboratory specialists [13] to move from service providers to leaders in patient assessment, helping to personalize diagnostics and treatments, improve the quality of digital data, and thereby support reductions in healthcare costs. Digital microbiology may also substantially impact public health and

* Corresponding author: A. Egli, Clinical Bacteriology and Mycology, University Hospital Basel, Petersgraben 4, 4031, Basel, Switzerland.

E-mail address: adrian.egli@usb.ch (A. Egli).

pathogen surveillance [14]. In order to enable digitalization, microbiology laboratories need to build a core expertise in digital medicine – including perception, know-how, and infrastructure on all aspects of data handling [15,16]. This review article aims to improve the general understanding of the most important aspects of digitalization, machine learning and artificial intelligence in the pre-to post-analytical process of clinical microbiology diagnostics.

Opportunities for digitalization in the microbiology diagnostic process

The diagnostic process in clinical microbiology is split into pre-analytical, analytical and post-analytical steps [17] and forms a circle of material and information flow. Table 1 highlights specific opportunities for digitalization in this process using the example of sepsis management.

Pre-analytics addresses the collection and quality of samples transported to the laboratory. For example, the filling volume of blood culture flasks, which directly correlates with positivity rates and the analytical sensitivity of blood culture diagnostics [18]. Modern blood culture systems provide automated weighting of blood cultures to determine the collected volumes and provide a feedback to the laboratory information system (LIS) [19]. Additional examples of pre-analytical quality include the detection of contaminated blood cultures due to skin flora such as *Staphylococcus epidermidis*, other coagulase negative staphylococci, and *Cutibacterium acnes*. Based on criteria of a systemic inflammatory response syndrome (SIRS) and the presence of a central venous line, the risk of blood culture contamination can be assessed [20]. In the future, the combination of LIS and electronic health record (EHR) data may allow more sophisticated feedback loops and provide automated quality assessments reports to the microbiologist and clinician.

Another important pre-analytical aspect is diagnostic stewardship. Diagnostic stewardship incorporates the concept of recommending the best diagnostic approach for a given situation [21–23]. Digital solutions in this field may range from digital twins [24,25] to machine-learning-based algorithms in smartphone apps [26] or chatbots [27,28]. Recently, chatbots have been developed to support the diagnostic evaluation and to recommend immediate measures, when patients are exposed to SARS-CoV-2 [27]. Similar to a microbiologist

consultant, a chatbot may provide helpful diagnostic information and advice, e.g., on the correct transport media for a sample, assay costs, the expected turn-around time, and test performance in specific sample types. Such an interactive tool may be a first source of information for routine and repetitive questions, and could support the pre-analytical quality management. In our vision, the digital twin works similarly to a smart shopping list, suggesting additional laboratory tests, which were previously ordered in the presence of similar patient characteristics. Thereby, such a tool may utilize the experience of other users. As an example, in a critically ill immunosuppressed patient with sepsis, a panel PCR directly from positive blood culture may speed up the species identification and result in an adaptation of the antibiotic treatment given [29], whereas in an otherwise healthy younger patient, standard culture based identification may be sufficient.

Test performance and data generation within the laboratory are parts of analytics. As an example, automated microscopy allows high-resolution images to be acquired of smears from positive blood cultures and can categorize Gram staining with high sensitivity and specificity [30,31]. Besides state-of-the-art automated microscopes, smartphones can also be used for image analysis of microscopy data [32,33]. Automated plate reading systems act similarly on pattern recognition and can reliably recognize bacterial growth on an agar plate and could be used to pre-screen culture plates [34–38]. Such automated plate reading systems are currently established in many European laboratories as part of the ongoing automation process. Reading of E-tests and inhibition zone diameters around antibiotic-impregnated discs can also be automatized with well-developed reading software [39,40]. Expert systems to interpret antimicrobial resistance profiles have already been in use for many years. Usually medical validation of phenotypic resistance profiles are performed to check whether there are unusual resistances, which prompts additional testing or confirmation, e.g., detection of a potential extended spectrum beta-lactamase (ESBL) or carbapenemase-producing bacteria due to suspicious MICs levels. This would trigger subsequent phenotypic or genotypic analysis [41,42]. Additional examples in the analytical step include the identification of bacterial and fungal species using matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometry. Machine learning based algorithms can link mass spectral profiles to specific clinical phenotypes such as antibiotic resistance [43,44].

Table 1
Aspects of digital microbiology in the diagnostic process

Process	Aspect	Example	References
Pre-analytics	Quality control	What is the sample quality? – Automated measurement and feedback regarding the correct filling of blood culture bottles. – Automated assessment of sample contamination including species and clinical score	[18–20]
	Diagnostic stewardship	Which additional diagnostic test should be ordered? – Suggestion based on a digital twin, smartphone app, or chatbot	[26–28]
Analytics	Quality control	How reliable is the analytical performance of a test? – Surveillance of reagent lots performance with internal and external controls and automated reporting in connection to specific used lots of time	[129]
	Imaging	Are there bacteria on the microscope slide? – Automated image acquisition with a microscope and scan for pathogen-like structures and category	[30,32,33]
	Plate reading	Is there bacterial growth on the plate? – Automated image acquisition and scan for colonies and subsequent identification (telebacteriology)	[36–38]
	Expert system	Does the detected resistance profile make sense? – Medical validation of antibiotic resistance profiles with expert database	[41,42]
	Public Health	Is there a potential outbreak? – Automated screening for pathogen similarities, e.g., resistance profile or automated bioinformatics	[130,131]
Post-analytics	Highlight important data	Is there a potential bacterial phenotype? – Detection of resistance by analysing MALDI-TOF spectra	[43,44]
	Sepsis treatment	What is the best treatment for the patient? – Prediction of sepsis, and best treatment, e.g., volume and antibiotics for the patient	[47–49]

Data visualization, communication, and clinical decision making are parts of post-analytics. Dashboards are an increasingly common way to visualize and summarize data. More complex applications include clinical decision support systems, e.g., for antibiotic stewardship. Empiric antibiotic treatment is dependent on knowledge of local antibiotic susceptibilities in each specific bacterial strain. Based on specific clinical information, such as patient demographics (age or gender) or specific wards, the empiric treatment may be further adapted [45]. In the near future, we may expect clinical decision support systems based on machine learning to provide automated feedback regarding empiric antibiotic prescription adapted to specific patient groups [46]. As a next step, more complex datasets will also be analysed. As physiology and laboratory parameters can rapidly change during an infection, time-series data greatly impact the predictive values of such algorithms – similar to a doctor, who observes the patient during disease progression – machine learning algorithms will also follow the patient's data stream. Recently, a series of studies has shown the impact of high-frequency physiological parameters in ICUs on the prediction of sepsis [47–49] or meningitis [50,51]. These studies are retrospective analyses and prospective controlled validation studies are largely missing in the field. Therefore, although our expectations for digital microbiology may be high, we should remain critical and carefully address the associated challenges.

Challenges of digitalization in the microbiology diagnostic process

The collection, quality control and cleaning, storage, security and protection, stewardship and governance, interoperability and interconnection, reporting and visualization, versioning, and sharing of data pose considerable challenges for big data in microbiology diagnostic laboratories. Some of these data handling aspects may be managed with a profound understanding of the laboratory and data workflows and clinical microbiology informatics [52]. However, rapidly developing computer technologies and increasing availability of storage space pose an important challenge itself for microbiologists and infectious disease experts: the amount of data with a deep medical context will explode over the next few years. Three trends currently explain this explosion of information: (a) larger number of fields are being collected, (b) the replacement of aggregate by person-specific data, and (c) the start of collecting new person-specific data [53]. In 2010, the global stored information amount already exceeded 1000 exabytes of data (i.e. $10e21$). Moreover, Densen and colleagues postulated a dramatic reduction in the half-life of life-science knowledge to only 73 days in 2020 [54]. In clinical microbiology laboratories, there is a similar exponential accumulation of routine data, e.g., MALDI-TOF mass spectra, photo-documented microscopy slides, pictures of agar plates (telebacteriology), sequencing data (microbial genomics, microbiota analysis), results of real-time PCR, and serological assays. Gigabytes of data are already produced every day and are stored for quality control, accreditation, legal reasons, and research.

Due to the increasing quantity of data (explosion of information), it will soon become almost impossible for a human to keep a clear view and interconnect the most important and relevant pieces of information [16]. Today, clinical colleagues have to access several computer programmes to collect information from various sources. The large amount of opaque data results in a demand to report the most critical results directly to the clinician, e.g., via phone calls of bacteraemia cases [55] [– thereby flagging most critical results. Digital tools will need to efficiently facilitate the raw-data-to-knowledge process [56]. Laboratory specialists, lab technicians, physicians, nurses, and information technology (IT) experts will clearly be challenged to handle this rapidly approaching information tsunami. New

communication and visualization strategies will be important and the interface between laboratory and clinics has to evolve and adapt. As examples, dashboards summarize the most critical clinical information and help to communicate complex data [57,58] or pop-up windows of automated alerting systems indicate critical results in specific patient groups [59] in a targeted fashion.

Data accumulation and complexity will further amplify, as we use more advanced technologies to achieve a detailed and structured description of the microbiological data (e.g., the Microbiology Investigation Criteria for Reporting Objectively (MICRO) criteria [60]). In clinical microbiology, the introduction of panel PCRs was only a first step. Molecular diagnostics will move towards metagenomics applications [61,62] in the next years. Thereby, the information will become more complex including pathogen and host genetic data. The problem is that (a) in non-primary sterile sites, multiple organisms can be detected – potential pathogens and colonizers – with sometimes unknown significance, and (b) not all antibiotic resistance genes can be linked to a specific species. For example, coagulase-negative staphylococci with oxacillin resistance in a sputum sample, along with *Staphylococcus aureus* may be misinterpreted as presence of methicillin-resistant *S. aureus* (MRSA) [63]. It will be crucial to identify those pathogens that are relevant and know which resistance mechanisms are linked to a specific pathogen. Simply providing an (endless) list of bacteria and resistance genes may result in non-reflected antibiotic usage and the treatment of a lab result and not of the patient. Future software algorithms could antibiotic resistance genes and Operational Taxonomic Units (OTUs) in bacterial networks of acute or chronic infectious diseases [64]. In addition, long-read sequencing metagenomics may overcome the problem of linking resistance genes to a specific species [65,66]. In this situation, machine learning could offer to analyse the complex interactions of bacterial networks within the host and help to better understand the data [67,68], for example understanding treatment failure by analysing mixed bacterial samples in the context of enzymatic inactivation by commensal species associated with a pathogen. Even more detailed data is expected from metabolomics, proteomic and transcriptomic analysis during infections in the next 10–15 years.

Besides the rapid increase of data and its associated problems, the changes anticipated with the new technologies may be very profound for laboratory personnel. Change management on various levels will be essential to manage expectations and fears linked to digitalization [69–71]. Whereas many classical tasks such as manual culture plate reading and microscopy may disappear over the next years, new aspects will fill these gaps for laboratory technicians and microbiologists including dry-lab tasks such as data handling and analysis for diagnostics, research and development. The educational portfolio of all laboratory personnel – clinical microbiologists and lab technicians – has to adapt to meet the new requirements of digital microbiology.

A first step: data structure and interoperability

Datasets collected in the clinical data warehouse will ideally allow more detailed analysis of infectious diseases (Fig. 1) [72–74]. Machine learning algorithms require large, structured, interoperable, and interconnected datasets. Healthcare data must be further standardized and annotated with internationally recognized definitions [75,76]. Ontologies help to structure data in such a way by using a common vocabulary, and allow the determination of relations of variables within a data model [77]. As an example, antibiotic susceptibility testing may be performed with various technical methods providing different sensitivities, error margins, and interpretation guidelines of breakpoints – the ontology term allows the specific description of the method in a machine-

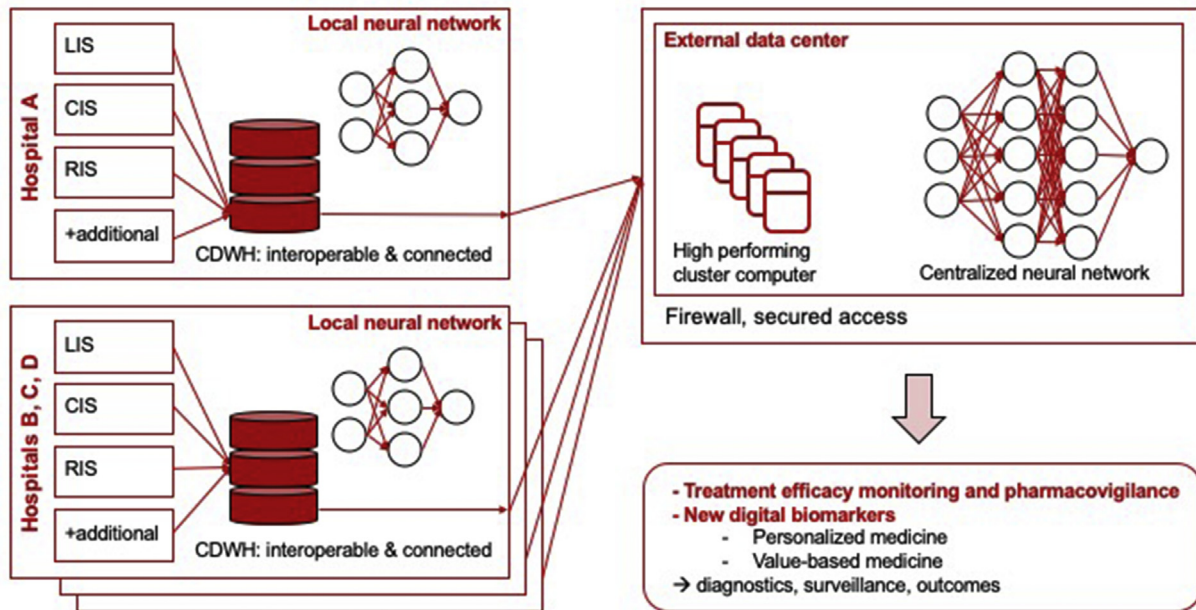


Fig. 1. Concept of data handling within and across institutions. Local data warehouses with local cluster computers transfer interconnected and interoperable data for diagnostics, research and development to larger clusters allowing the enrichment of datasets. Clinical Data Warehouse (CDWH), Clinical Information System (CIS), Laboratory Information System (LIS), Radiology Information System (RIS).

readable format and helps to compare results across different datasets. Various ontologies exist for clinical, laboratory and microbiological data such as the WHO's International Classification of Diseases (ICD), Systematized Nomenclature of Human and Veterinary Medicine Clinical Terms (SNOMED CT; <http://www.snomed.org/>; [78]), Logical Observation Identifiers Names and Codes (LOINC; <https://loinc.org/>; [79]), or the Integrated Rapid Infectious Diseases Analysis (IRIDA; <https://www.irida.ca/>).

Besides the clear requirements for structure and interoperability of data, also data security and protection, and the versioning of datasets are important. Sensitive healthcare data should only be transferred if anonymized or encoded and simultaneously encrypted [80,81]. For this, specific data security standards and scripts are necessary [82–84]. Data safety breaches may have severe consequences, with more than 70% of recent hospital data breaches including sensitive demographic or financial info that could lead to identity theft [85]. For certain databases, the blockchain technology provides interesting solutions regarding data safety and could be particularly well suited to public health surveillance or clinical trial management [86–88].

An underestimated challenge is the versioning of ontologies, guidelines, and recommendations. Maintenance and curation of databases are costly, but this remains a highly critical element directly linked to the quality of a database [89]. As an example, with every annual updated antibiotic resistance interpretation by the European Committee on Antimicrobial Susceptibility Testing (EUCAST) or Clinical and Laboratory Standards Institute (CLSI) antibiotic breakpoints may change and comparability across years is jeopardized [90]. The new breakpoints in EUCAST v10 for *Pseudomonas aeruginosa* resulted in an increased rate of the intermediate category for most penicillin antibiotics, but at the same time, the clinical meaning of the intermediate category was also changed (www.eucast.org). Comparing only the categorical trends, without further knowledge of the version used, harbours the risk for false interpretations. Therefore, changes in databases must be well documented and tracked, otherwise temporal trends cannot be reliably analysed. A way around extensive versioning may be the storage of raw data. In the given example, this would be the storage

of minimal inhibitory concentrations, which could be re-used using different breakpoints. Storage of raw data also has specific challenges such as storage space, changes in data formats, and can be more demanding regarding data protection.

The previously mentioned concepts for data handling have been used for a series of large healthcare data repositories, e.g., the Medical Information Mart for Intensive Care (MIMIC)-III (<https://mimic.physionet.org/>; [91]) dataset or the eICU collaborative research database (<https://eicu-crd.mit.edu/>; [92]). MIMIC-III and eICU are large databases supporting sepsis research. Similarly, the Swiss Personalized Health Network (SPHN; www.sphn.ch) currently supports digitalization projects throughout Switzerland enabling a national data infrastructure, ensuring data interoperability of local and regional information systems, with special emphasis on clinical data management systems allowing effective exchange of patient data. The SPHN driver project 'Personalized Swiss Sepsis Study' integrates data from clinical microbiology, infectious diseases and intensive care medicine of all University Hospitals. The goal is to discover digital biomarkers for early sepsis recognition and prediction of mortality using machine learning algorithms (www.sphn.ch).

Epidemiological databases can also benefit from structured data. For example, Pulsenet is a large sequencing repository with its main focus on food-borne pathogens (www.cdc.gov/pulsenet; [93]). Other platforms such as microreact ([94]) or nextstrain (<https://nextstrain.org/>; [95]) visualize sequencing data either on a project basis or as semi-automated surveillance tools, which access public sources such as GISAID (<https://www.gisaid.org/>) in the case of influenza or SARS-CoV-2. Similarly, the Swiss Pathogen Surveillance Platform (www.spsp.ch), aims to establish one health network for sharing of whole genome sequencing (WGS) data of pathogens for public health surveillance and epidemiological research [96]. Within the SPSP platform, demographic, epidemiological and microbiological metadata is interconnected and interoperable to add spatio-temporal, clinical and veterinary contexts. A prototype is currently being tested for the transmission of methicillin-resistant *S. aureus* between veterinary and human sources. Additional databases allow the exploration and cross-

analysis of host-associated data in infectious diseases [97,98] or the environment integrating the previously mentioned one health approach, e.g., the Earth Microbiome Project [99] or the China National Gene Bank (db.cngb.org).

A second step: legal framework

Data collection, analysis and exchange must follow legal and regulatory requirements. Therefore, an ethical evaluation is mandatory as well as a patient consent, e.g., via a study-specific or general consent [100–102]. At present, general consents are not authorized in some countries to prevent further data usage in other studies than that explained to the patient. The evolution of these ethical rules appears mandatory as collected data will not become exploitable, and this will probably raise numerous additional ethical issues as well. Shared datasets can be tremendously useful to improve, e.g., public health surveillance [103–107] and not sharing the data in emergencies may be unethical as well [96]. Globally, different data protection, human research and epidemiological laws exist. In Europe, the General Data Protection Regulation (GDPR) 2016/679 is enforced and has to be followed for European citizens (eur-lex.europa.eu). Non-genetic anonymized data is often excluded from regulation. However, there is an ongoing debate as to what anonymization means. The interface of research/surveillance and data protection generates additional challenges [108]. If larger datasets are shared between centres, ethical committees usually ask for a detailed data-management plan as part of a data transfer and use agreement (DTUA). In such a context, it is also often advisable to generate a collaboration or consortium agreement (CA) between research institutions, regulating the way of data collection, storage, access rights and protection, duties, responsibilities, publications, and intellectual properties.

In research, there is an increasing trend for data sharing [107,109,110]. Whereas traditionally, research groups were silos of innovation and technologies, nowadays cutting-edge research often happens in international teams. Across institutions, a framework should be generated enabling research with pragmatic solutions which will help patients, physicians and society. Data sharing allows us to drive innovation. In public funded projects the Findability, Accessibility, Interoperability and Reusability (FAIR) principles (<https://libereurope.eu>) are often used and multiple scientific journals follow these important guidelines. These principles cover: (a) data and supplementary materials having sufficiently rich metadata and a unique and persistent identifier, (b) metadata and data being understandable to humans and machines, (c) data is deposited in a trusted repository, (d) metadata using a formal, accessible, shared, and broadly applicable language for knowledge representation, and (e) data and collections having a clear usage licenses and provide accurate information on their provenance [111].

Ways to applications: use of machine learning in the modern microbiology laboratory

Machine learning is based on sample data (a training or discovery dataset) in order to make predictions or decisions without being explicitly programmed to perform that task [9,112]. Machine learning algorithms may be used at each step of the microbiological diagnostic process from pre- to post-analytics, helping us to deal with the increasing quantities and complexity of data [113,114] (Table 1). Human analytical capacity has reached its limits to (a) grasp the huge amount of available complex data, (b) interconnect data in single patients, groups and across the population, and (c) draw meaningful conclusions from this. Machine learning

algorithms can overcome these limits, by using structured data and by helping to recognize patterns with supervised and unsupervised methods [115–117].

Besides the diagnostic process, a series of research studies have been performed focusing on machine learning of infectious diseases: prediction of infection on hospital admission [118], detection of urinary tract infections [119], self-reported influenza-like illness [120], prediction of complications in *Clostridioides difficile* infection [121], identification of antibiotic drug resistance in *Mycobacterium tuberculosis* [122], detection of ventilator-associated pneumonia with *P. aeruginosa* on intensive care units [123], estimation of outcomes of shigellosis [124], drug discovery for new antibiotics [125], prediction of side effects [126,127], and pharmacokinetic/pharmacodynamic (PK/PD) models of antibiotics [128], and many more.

The establishment of machine learning algorithms in microbiological routine workflows requires a profound system understanding of the pre- to post-analytical steps and data handling knowledge. Where are the most important interfaces in the workflow? What are the current gaps in communication? Where and how could the quality of the process(es) be improved with digital technologies? The answers require a local in-depth analysis of the diagnostic process and digital environment. The development of digital microbiology should be closely monitored by the microbiologist as (a) understanding and access of the pre-analytical, analytical and post-analytical process management is key, (b) data handling is easiest at the point where the data is actually produced, and (c) laboratory personnel is familiar with standardization and regulatory aspects of diagnostic tests. In general, incentives are needed to further support all aspects of data handling in laboratory medicine – including standardization data structures and machine learning algorithms.

Conclusion

Digitalization in healthcare already shows a profound impact on patients. It is expected that the developments started will further gain momentum. Machine learning radically changes the way we handle healthcare-related data – including data on clinical microbiology and infectious diseases. We will probably move from the internet-of-things environment (interconnected datasets in a patient with a disease) to the internet-of-bodies with (implanted) devices, providing detailed healthcare data also in a disease-free time. In addition, developments of molecular diagnostics such as metagenomics will increase the data complexity. Current trends indicate that the importance of laboratory diagnostics will further increase over the next decade. This means that the clinical microbiologist of today needs to (a) get familiar with the concepts of digital microbiology, (ii) get educated on data handling and (iii) anticipate the low hanging fruits such as microbiology dashboards, expert systems, and image analysis of microscopy slides and plate reading. Now is the time for clinical microbiology laboratories to evaluate their data handling processes and available infrastructures, including storage and data transfer workflows. We need to develop strategies for the next 5–10 years to face the opportunities and challenges ahead of us. Our community should anticipate the advances in digitalization and develop concepts including machine-readable formats and interoperability across centres to further improve patient care.

Transparency declaration

None of the authors has any conflicts of interest. The article was supported by grants of the Swiss Personalized Health Network driver project on Sepsis (www.sphn.ch) and the Swiss National

Science Foundation, NRP72 program (407240_177504) to all authors.

Author contributions

A.E. wrote the original draft; A.E., J.S. and G.G. reviewed and edited the manuscript.

Acknowledgements

The authors are grateful to Dr. Helena Seth-Smith, Dr. Kirstine K. Sogaard and Dr. Vladimira Hinic (University Hospital Basel) for critical feedback regarding the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cmi.2020.06.023>.

References

- Wilson N, MacDonald EJ, Mansoor OD, Morgan J. In bed with Siri and Google Assistant: a comparison of sexual health advice. *BMJ* 2017;359:j5635. <https://doi.org/10.1136/bmj.j5635>.
- F, M. B., et al. Disruptive technologies for environment and health research: an overview of artificial intelligence, Blockchain, and Internet of Things. *Int J Environ Res Public Health* 2019;16:3847. <https://doi.org/10.3390/ijerph16203847>.
- Baig MM, Afifi S, Gholamhosseini H, Mirza F. A systematic review of wearable sensors and IoT-based monitoring applications for older adults – a focus on ageing population and independent living. *J Med Syst* 2019;43:233. <https://doi.org/10.1007/s10916-019-1365-7>.
- Tom-Aba D, Nguku PM, Arinze CC, Krause G. Assessing the concepts and designs of 58 mobile apps for the management of the 2014–2015 West Africa Ebola outbreak: systematic review. *JMIR Public Health Surveill* 2018;4:e68. <https://doi.org/10.2196/publichealth.9015>.
- Amukele T. Current state of drones in healthcare: challenges and opportunities. *J Appl Lab Med* 2019;4:296–8. <https://doi.org/10.1373/jalm.2019.030106>.
- Poon NC, Sung JJ. Self-driving cars and AI-assisted endoscopy: who should take the responsibility when things go wrong? *J Gastroenterol Hepatol* 2019;34:625–6. <https://doi.org/10.1111/jgh.14641>.
- Comellas MJ, et al. Evaluation of a new digital automated glycemic pattern detection tool. *Diabetes Technol Ther* 2017;19:633–40. <https://doi.org/10.1089/dia.2017.0180>.
- Velardo C, et al. Digital health system for personalised COPD long-term management. *BMC Med Inform Decis Mak* 2017;17:19. <https://doi.org/10.1186/s12911-017-0414-8>.
- Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019;25:44–56. <https://doi.org/10.1038/s41591-018-0300-7>.
- Bhavnani SP, Narula J, Sengupta PP. Mobile technology and the digitization of healthcare. *Eur Heart J* 2016;37:1428–38. <https://doi.org/10.1093/eurheartj/ehv770>.
- Sanchez-Pinto LN, Luo Y, Churpek MM. Big data and data science in critical care. *Chest* 2018;154:1239–48. <https://doi.org/10.1016/j.chest.2018.04.037>.
- Dolen V, et al. In: *Changing diagnostic paradigms for microbiology: report on an American academy of microbiology colloquium held in Washington, DC, from 17 to 18 October 2016*; 2017.
- 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. <https://doi.org/10.1016/j.clinbiochem.2019.04.013>.
- Basak SC, et al. Computer-assisted and data driven approaches for surveillance, drug discovery, and vaccine design for the Zika virus. *Pharmaceuticals (Basel)* 2019;12. <https://doi.org/10.3390/ph12040157>.
- Cosgriff CV, Celi LA, Stone DJ. Critical care, critical data. *Biomed Eng Comput Biol* 2019;10. <https://doi.org/10.1177/1179597219856564>.
- Galetsi P, Katsaliaki K, Kumar S. Values, challenges and future directions of big data analytics in healthcare: a systematic review. *Soc Sci Med* 2019;241:112533. <https://doi.org/10.1016/j.socscimed.2019.112533>.
- Chan JZ, Pallen MJ, Oppenheim B, Constantinidou C. Genome sequencing in clinical microbiology. *Nat Biotechnol* 2012;30:1068–71. <https://doi.org/10.1038/nbt.2410>.
- Henning C, Aygul N, Dinnetz P, Wallgren K, Ozenci V. Detailed analysis of the characteristics of sample volume in blood culture bottles. *J Clin Microbiol* 2019;57. <https://doi.org/10.1128/JCM.00268-19>.
- Khare R, et al. Active monitoring and feedback to improve blood culture fill volumes and positivity across a large integrated health system. *Clin Infect Dis* 2020;70:262–8. <https://doi.org/10.1093/cid/ciz198>.
- Elzi L, et al. How to discriminate contamination from bloodstream infection due to coagulase-negative staphylococci: a prospective study with 654 patients. *Clin Microbiol Infect* 2012;18:E355–61. <https://doi.org/10.1111/j.1469-0691.2012.03964.x>.
- Bou G, Canton R, Martinez-Martinez L, Navarro D, Vila J. Fundamentals and implementation of microbiological diagnostic stewardship programs. *Enferm Infecc Microbiol Clin* 2020. <https://doi.org/10.1016/j.eimc.2020.02.019>.
- Broadhurst MJ, et al. Utilization, yield, and accuracy of the FilmArray Meningitis/Encephalitis panel with diagnostic stewardship and testing algorithm. *J Clin Microbiol* 2020. <https://doi.org/10.1128/JCM.00311-20>.
- Poelman R, et al. Improved diagnostic policy for respiratory tract infections essential for patient management in the emergency department. *Future Microbiol* 2020. <https://doi.org/10.2217/fmb-2019-0119>.
- Bruynseels K, Santoni de Sio F, van den Hoven J. Digital twins in health care: ethical implications of an emerging engineering paradigm. *Front Genet* 2018;9:31. <https://doi.org/10.3389/fgene.2018.00031>.
- Bjornsson B, et al. Digital twins to personalize medicine. *Genome Med* 2019;12:4. <https://doi.org/10.1186/s13073-019-0701-3>.
- Piau A, Crissey R, Brechemier D, Balardy L, Nourhashemi F. A smartphone Chatbot application to optimize monitoring of older patients with cancer. *Int J Med Inform* 2019;128:18–23. <https://doi.org/10.1016/j.ijmedinf.2019.05.013>.
- Battineni G, Chintalapudi N, Amenta F. AI Chatbot design during an epidemic like the novel Coronavirus. *Healthcare (Basel)* 2020;8. <https://doi.org/10.3390/healthcare8020154>.
- Zand A, et al. An exploration into the use of a chatbot for patients with inflammatory bowel diseases: retrospective cohort study. *J Med Internet Res* 2020;22:e15589. <https://doi.org/10.2196/15589>.
- Verroken A, Despas N, Rodriguez-Villalobos H, Lateral PF. The impact of a rapid molecular identification test on positive blood cultures from critically ill with bacteremia: a pre-post intervention study. *PLoS One* 2019;14:e0223122. <https://doi.org/10.1371/journal.pone.0223122>.
- 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>.
- 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. <https://doi.org/10.1128/JCM.01521-17>.
- Linares M, et al. Collaborative intelligence and gamification for on-line malaria species differentiation. *Malar J* 2019;18:21. <https://doi.org/10.1186/s12936-019-2662-9>.
- Perkel JM. Pocket laboratories. *Nature* 2017;545:119–21. <https://doi.org/10.1038/545119a>.
- Croxatto A, Prod'homme G, Faverjon F, Rochais Y, Greub G. Laboratory automation in clinical bacteriology: what system to choose? *Clin Microbiol Infect* 2016;22:217–35. <https://doi.org/10.1016/j.cmi.2015.09.030>.
- 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. <https://doi.org/10.1016/j.bj.2017.09.001>.
- Faron ML, Buchan BW, Samra H, Ledebor NA. Evaluation of WASPLab software to automatically read chromID CPS elite agar for reporting of urine cultures. *J Clin Microbiol* 2019;58. <https://doi.org/10.1128/JCM.00540-19>.
- Glasson J, et al. Multicenter evaluation of an image analysis device (APAS): comparison between digital image and traditional plate reading using urine cultures. *Ann Lab Med* 2017;37:499–504. <https://doi.org/10.3343/alm.2017.37.6.499>.
- 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. <https://doi.org/10.1128/JCM.00811-19>.
- Smith KP, Richmond DL, Brennan-Krohn T, Elliott HL, Kirby JE. Development of mast: a microscopy-based antimicrobial susceptibility testing platform. *SLAS Technol* 2017;22:662–74. <https://doi.org/10.1177/2472630317727721>.
- Strauss M, Zoabi K, Sagas D, Reznik-Gitlitz B, Colodner R. Evaluation of Bio-Rad(R) discs for antimicrobial susceptibility testing by disc diffusion and the ADAGIO system for the automatic reading and interpretation of results. *Eur J Clin Microbiol Infect Dis* 2020;39:375–84. <https://doi.org/10.1007/s10096-019-03735-4>.
- Karagoz A, Acar S, Korkoca H. Characterization of *Klebsiella* isolates by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) and determination of antimicrobial resistance with VITEK 2 advanced expert system (AES). *Turk J Med Sci* 2015;45:1335–44. <https://doi.org/10.3906/sag-1401-32>.
- Winstanley T, Courvalin P. Expert systems in clinical microbiology. *Clin Microbiol Rev* 2011;24:515–56. <https://doi.org/10.1128/CMR.00061-10>.
- Sousa T, et al. Putative protein biomarkers of *Escherichia coli* antibiotic multiresistance identified by MALDI mass spectrometry. *Biology (Basel)* 2020;9. <https://doi.org/10.3390/biology9030056>.
- 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>.

- [45] Bielicki JA, et al. Selecting appropriate empirical antibiotic regimens for paediatric bloodstream infections: application of a Bayesian decision model to local and pooled antimicrobial resistance surveillance data. *J Antimicrob Chemother* 2016;71:794–802. <https://doi.org/10.1093/jac/dkv397>.
- [46] Beaudoin M, Kabanza F, Nault V, Valiquette L. Evaluation of a machine learning capability for a clinical decision support system to enhance antimicrobial stewardship programs. *Artif Intell Med* 2016;68:29–36. <https://doi.org/10.1016/j.artmed.2016.02.001>.
- [47] Komorowski M. Clinical management of sepsis can be improved by artificial intelligence: yes. *Intensive Care Med* 2020;46:375–7. <https://doi.org/10.1007/s00134-019-05898-2>.
- [48] Schinkel M, Paranjape K, Nannan Panday RS, Skyttberg N, Nanayakkara PWB. Clinical applications of artificial intelligence in sepsis: a narrative review. *Comput Biol Med* 2019;115:103488. <https://doi.org/10.1016/j.combiomed.2019.103488>.
- [49] van Wyk F, et al. A minimal set of physiometers in continuous high frequency data streams predict adult sepsis onset earlier. *Int J Med Inform* 2019;122:55–62. <https://doi.org/10.1016/j.ijmedinf.2018.12.002>.
- [50] Mago VK, Mehta R, Woolrych R, Papageorgiou EI. Supporting meningitis diagnosis amongst infants and children through the use of fuzzy cognitive mapping. *BMC Med Inform Decis Mak* 2012;12:98. <https://doi.org/10.1186/1472-6947-12-98>.
- [51] Savin I, et al. Healthcare-associated ventriculitis and meningitis in a neuro-ICU: incidence and risk factors selected by machine learning approach. *J Crit Care* 2018;45:95–104. <https://doi.org/10.1016/j.jcrc.2018.01.022>.
- [52] Rhoads DD, Sintchenko V, Rauch CA, Pantanowitz L. Clinical microbiology informatics. *Clin Microbiol Rev* 2014;27:1025–47. <https://doi.org/10.1128/CMR.00049-14>.
- [53] Sweeney L. In: Zayatz L, Doyle P, Theeuwes J, Lane J, editors. Confidentiality, disclosure, and data access: theory and practical applications for statistical agencies. Urban Institute; 2001. p. 1–26.
- [54] Densen P. Challenges and opportunities facing medical education. *Trans Am Clin Climatol Assoc* 2011;122:48–58.
- [55] Bunsow E, et al. Improved sepsis alert with a telephone call from the clinical microbiology laboratory: a clinical trial. *Medicine (Baltimore)* 2015;94:e1454. <https://doi.org/10.1097/MD.0000000000001454>.
- [56] Wang X, Williams C, Liu ZH, Croghan J. Big data management challenges in health research—a literature review. *Brief Bioinform* 2019;20:156–67. <https://doi.org/10.1093/bib/bbx086>.
- [57] Meng Y, et al. Lessons learned in the development of a web-based surveillance reporting system and dashboard to monitor acute febrile illnesses in Guangdong and Yunnan Provinces, China, 2017–2019. *Health Secur* 2020;18:S14–22. <https://doi.org/10.1089/hs.2019.0079>.
- [58] Raban MS, Bamford C, Joalay Y, Harrison MC. Impact of an educational intervention and clinical performance dashboard on neonatal bloodstream infections. *S Afr Med J* 2015;105:564–6. <https://doi.org/10.7196/SAMJnew.7764>.
- [59] Buckley MS, et al. Trigger alerts associated with laboratory abnormalities on identifying potentially preventable adverse drug events in the intensive care unit and general ward. *Ther Adv Drug Saf* 2018;9:207–17. <https://doi.org/10.1177/2042098618760995>.
- [60] Turner P, et al. Microbiology investigation criteria for reporting objectively (MICRO): a framework for the reporting and interpretation of clinical microbiology data. *BMC Med* 2019;17:70. <https://doi.org/10.1186/s12916-019-1301-1>.
- [61] Chiu CY, Miller SA. Clinical metagenomics. *Nat Rev Genet* 2019;20:341–55. <https://doi.org/10.1038/s41576-019-0113-7>.
- [62] Miller S, Chiu C, Rodino KG, Miller MB. Should we be performing metagenomic next-generation sequencing for infectious disease diagnosis in the clinical laboratory? *J Clin Microbiol* 2019. <https://doi.org/10.1128/JCM.01739-19>.
- [63] Becker K, et al. Does nasal cocolonization by methicillin-resistant coagulase-negative staphylococci and methicillin-susceptible *Staphylococcus aureus* strains occur frequently enough to represent a risk of false-positive methicillin-resistant *S. aureus* determinations by molecular methods? *J Clin Microbiol* 2006;44:229–31. <https://doi.org/10.1128/JCM.44.1.229-231.2006>.
- [64] Brito IL, Alm EJ. Tracking strains in the microbiome: insights from metagenomics and models. *Front Microbiol* 2016;7:712. <https://doi.org/10.3389/fmicb.2016.00712>.
- [65] Bertrand D, et al. Hybrid metagenomic assembly enables high-resolution analysis of resistance determinants and mobile elements in human microbiomes. *Nat Biotechnol* 2019;37:937–44. <https://doi.org/10.1038/s41587-019-0191-2>.
- [66] Wood RL, et al. Analysis of identification method for bacterial species and antibiotic resistance genes using optical data from DNA oligomers. *Front Microbiol* 2020;11:257. <https://doi.org/10.3389/fmicb.2020.00257>.
- [67] Mahnic A, et al. Distinct types of gut microbiota dysbiosis in hospitalized gastroenterological patients are disease non-related and characterized with the predominance of either Enterobacteriaceae or Enterococcus. *Front Microbiol* 2020;11:120. <https://doi.org/10.3389/fmicb.2020.00120>.
- [68] Zhou YH, Gallins P. A review and tutorial of machine learning methods for microbiome host trait prediction. *Front Genet* 2019;10:579. <https://doi.org/10.3389/fgene.2019.00579>.
- [69] Croxatto A, Greub G. Project management: importance for diagnostic laboratories. *Clin Microbiol Infect* 2017;23:434–40. <https://doi.org/10.1016/j.cmi.2017.04.007>.
- [70] Wiljer D, Hakim Z. Developing an artificial intelligence-enabled health care practice: rewiring health care professions for better care. *J Med Imaging Radiat Sci* 2019;50:S8–14. <https://doi.org/10.1016/j.jmir.2019.09.010>.
- [71] Zeike S, Choi KE, Lindert L, Pfaff H. Managers' well-being in the digital era: is it associated with perceived choice overload and pressure from digitalization? An exploratory study. *Int J Environ Res Public Health* 2019;16:1746. <https://doi.org/10.3390/ijerph16101746>.
- [72] Bouzille G, et al. Leveraging hospital big data to monitor flu epidemics. *Comput Methods Programs Biomed* 2018;154:153–60. <https://doi.org/10.1016/j.cmpb.2017.11.012>.
- [73] Ciofi Degli Atti ML, Pecoraro F, Piga S, Luzi D, Raponi M. Developing a surgical site infection surveillance system based on hospital unstructured clinical notes and text mining. *Surg Infect (Larchmt)* 2020. <https://doi.org/10.1089/sur.2019.238>.
- [74] Grammatico-Guillon L, et al. Antibiotic prescribing in outpatient children: a cohort from a clinical data warehouse. *Clin Pediatr (Phila)* 2019;58:681–90. <https://doi.org/10.1177/0009922819834278>.
- [75] Gordon CL, et al. Design and evaluation of a bacterial clinical infectious diseases ontology. *AMIA Annu Symp Proc* 2013:502–11. 2013.
- [76] Smith B, Scheuermann RH. Ontologies for clinical and translational research: Introduction. *J Biomed Inform* 2011;44:3–7. <https://doi.org/10.1016/j.jbi.2011.01.002>.
- [77] 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. <https://doi.org/10.1007/s10096-019-03501-6>.
- [78] Millar J. The need for a global language – SNOMED CT introduction. *Stud Health Technol Inform* 2016;225:683–5.
- [79] Tsao HM, et al. Toward automatic reporting of infectious diseases. *Stud Health Technol Inform* 2017;245:808–12.
- [80] Neame R. Effective sharing of health records, maintaining privacy: a practical schema. *Online J Public Health Inform* 2013;5:217. <https://doi.org/10.5210/ojphi.v5i2.4344>.
- [81] Renardi MB, Basjaruddin NC, Rakhman E. Securing electronic medical record in near field communication using advanced encryption standard (AES). *Technol Health Care* 2018;26:357–62. <https://doi.org/10.3233/THC-171140>.
- [82] Desjardins B, et al. DICOM images have been hacked! Now what? *AJR Am J Roentgenol* 2019;1–9. <https://doi.org/10.2214/AJR.19.21958>.
- [83] Raisaro JL, et al. Feasibility of homomorphic encryption for sharing I2B2 aggregate-level data in the cloud. *AMIA Jt Summits Transl Sci Proc* 2017:176–85. 2018.
- [84] Raisaro JL, et al. MedCo: enabling secure and privacy-preserving exploration of distributed clinical and genomic data. *IEEE/ACM Trans Comput Biol Bioinform* 2019;16:1328–41. <https://doi.org/10.1109/TCBB.2018.2854776>.
- [85] Jiang JX, Bai G. Types of information compromised in breaches of protected health information. *Ann Intern Med* 2019;172:159–60. <https://doi.org/10.7326/M19-1759>.
- [86] Agbo CC, Mahmoud QH, Eklund JM. Blockchain technology in healthcare: a systematic review. *Healthcare (Basel)* 2019;7:56. <https://doi.org/10.3390/healthcare7020056>.
- [87] Bhattacharya S, Singh A, Hossain MM. Strengthening public health surveillance through blockchain technology. *AIMS Public Health* 2019;6:326–33. <https://doi.org/10.3934/publichealth.2019.3.326>.
- [88] Wong DR, Bhattacharya S, Butte AJ. Prototype of running clinical trials in an untrustworthy environment using blockchain. *Nat Commun* 2019;10:917. <https://doi.org/10.1038/s41467-019-08874-y>.
- [89] Carle F, et al. Quality assessment of healthcare databases. *Epidemiol Bioestat Public Health* 2017;14. e12901-12901-12911.
- [90] Bork JT, et al. Impact of CLSI and EUCAST Cefepime breakpoint changes on the susceptibility reporting for Enterobacteriaceae. *Diagn Microbiol Infect Dis* 2017;89:328–33. <https://doi.org/10.1016/j.diagmicrobio.2017.08.020>.
- [91] Johnson AE, et al. MIMIC-III, a freely accessible critical care database. *Sci Data* 2016;3:160035. <https://doi.org/10.1038/sdata.2016.35>.
- [92] Pollard TJ, et al. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci Data* 2018;5:180178. <https://doi.org/10.1038/sdata.2018.178>.
- [93] Kubota KA, et al. PulseNet and the changing paradigm of laboratory-based surveillance for foodborne diseases. *Public Health Rep* 2019;134:225–85. <https://doi.org/10.1177/0033354919881650>.
- [94] Argimon S, et al. Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microb Genom* 2016;2:e000093. <https://doi.org/10.1099/mgen.0.000093>.
- [95] Hadfield J, et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 2018;34:4121–3. <https://doi.org/10.1093/bioinformatics/bty407>.
- [96] Egli A, et al. Improving the quality and workflow of bacterial genome sequencing and analysis: paving the way for a Switzerland-wide molecular epidemiological surveillance platform. *Swiss Med Wkly* 2018;148:w14693. <https://doi.org/10.4414/smw.2018.14693>.
- [97] Aevermann BD, et al. A comprehensive collection of systems biology data characterizing the host response to viral infection. *Sci Data* 2014;1:140033. <https://doi.org/10.1038/sdata.2014.33>.
- [98] Tan SL, Ganji G, Paepfer B, Proll S, Katze MG. Systems biology and the host response to viral infection. *Nat Biotechnol* 2007;25:1383–9. <https://doi.org/10.1038/nbt1207-1383>.

- [99] Thompson LR, et al. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 2017;551:457–63. <https://doi.org/10.1038/nature24621>.
- [100] Garattini C, Raffle J, Aisyah DN, Sartain F, Kozlakidis Z. Big data analytics, infectious diseases and associated ethical impacts. *Philos Technol* 2019;32: 69–85. <https://doi.org/10.1007/s13347-017-0278-y>.
- [101] Organizing Committee of the Madrid 2017 Critical Care Datathon, et al. Big data and machine learning in critical care: opportunities for collaborative research. *Med Intensiva* 2019;43:52–7. <https://doi.org/10.1016/j.medin.2018.06.002>.
- [102] Price 2nd WN, Cohen IG. Privacy in the age of medical big data. *Nat Med* 2019;25:37–43. <https://doi.org/10.1038/s41591-018-0272-7>.
- [103] Amid C, et al. The COMPARE data Hubs. *Database (oxford)* 2019, baz136. 2019. <https://doi.org/10.1093/database/baz136>.
- [104] Jackson C, Gardy JL, Shadiloo HC, Silva DS. Trust and the ethical challenges in the use of whole genome sequencing for tuberculosis surveillance: a qualitative study of stakeholder perspectives. *BMC Med Ethics* 2019;20:43. <https://doi.org/10.1186/s12910-019-0380-z>.
- [105] Saxena A, et al. Ethics preparedness: facilitating ethics review during outbreaks - recommendations from an expert panel. *BMC Med Ethics* 2019;20: 29. <https://doi.org/10.1186/s12910-019-0366-x>.
- [106] Caugant DA, Brynildsrud OB. *Neisseria meningitidis*: using genomics to understand diversity, evolution and pathogenesis. *Nat Rev Microbiol* 2020;18: 84–96. <https://doi.org/10.1038/s41579-019-0282-6>. 2019.
- [107] Littler K, et al. Progress in promoting data sharing in public health emergencies. *Bull World Health Organ* 2017;95:243. <https://doi.org/10.2471/BLT.17.192096>.
- [108] Kostkova P. Disease surveillance data sharing for public health: the next ethical frontiers. *Life Sci Soc Policy* 2018;14:16. <https://doi.org/10.1186/s40504-018-0078-x>.
- [109] Casey C, Li J, Berry M. Interorganizational collaboration in public health data sharing. *J Health Organ Manag* 2016;30:855–71. <https://doi.org/10.1108/JHOM-05-2015-0082>.
- [110] Staley J, et al. Novel data sharing agreement to accelerate big data translational research projects in the one health sphere. *Top Companion Anim Med* 2019;37:100367. <https://doi.org/10.1016/j.tcam.2019.100367>.
- [111] Brall C, Schroder-Back P, Maeckelberghe E. Ethical aspects of digital health from a justice point of view. *Eur J Public Health* 2019;29:18–22. <https://doi.org/10.1093/eurpub/ckz167>.
- [112] Bastanlar Y, Ozuysal M. Introduction to machine learning. *Methods Mol Biol* 2014;1107:105–28. https://doi.org/10.1007/978-1-62703-748-8_7.
- [113] 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>.
- [114] Peiffer-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>.
- [115] Macesic N, Polubriaginof F, Tatonetti NP. Machine learning: novel bioinformatics approaches for combating antimicrobial resistance. *Curr Opin Infect Dis* 2017;30:511–7. <https://doi.org/10.1097/QCO.0000000000000406>.
- [116] Peiffer-Smadja N, Rawson TM, Ahmad R, Buchard A, Pantelis G, Lescure F-X, et al. Machine learning for clinical decision support in infectious diseases: a narrative review of current applications. *Clin Microbiol Infect* 2020;26: 584–95. <https://doi.org/10.1016/j.cmi.2019.09.009>.
- [117] Wiens J, Shenoy ES. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. *Clin Infect Dis* 2018;66:149–53. <https://doi.org/10.1093/cid/cix731>.
- [118] Rawson TM, et al. Supervised machine learning for the prediction of infection on admission to hospital: a prospective observational cohort study. *J Antimicrob Chemother* 2019;74:1108–15. <https://doi.org/10.1093/jac/dky514>.
- [119] Burton RJ, Albur M, Eberl M, Cuff SM. Using artificial intelligence to reduce diagnostic workload without compromising detection of urinary tract infections. *BMC Med Inform Decis Mak* 2019;19:171. <https://doi.org/10.1186/s12911-019-0878-9>.
- [120] Kalimeri K, et al. Unsupervised extraction of epidemic syndromes from participatory influenza surveillance self-reported symptoms. *PLoS Comput Biol* 2019;15:e1006173. <https://doi.org/10.1371/journal.pcbi.1006173>.
- [121] Li BY, Oh J, Young VB, Rao K, Wiens J. Using machine learning and the electronic health record to predict complicated *Clostridium difficile* infection. *Open Forum Infect Dis* 2019;6:ofz186. <https://doi.org/10.1093/ofid/ofz186>.
- [122] Chen ML, et al. Beyond multidrug resistance: leveraging rare variants with machine and statistical learning models in *Mycobacterium tuberculosis* resistance prediction. *EBioMedicine* 2019;43:356–69. <https://doi.org/10.1016/j.ebiom.2019.04.016>.
- [123] Liao YH, et al. Machine learning methods applied to predict ventilator-associated pneumonia with *Pseudomonas aeruginosa* infection via sensor array of electronic nose in intensive care unit. *Sensors (Basel)* 2019;19:1866. <https://doi.org/10.3390/s19081866>.
- [124] Adamker G, et al. Prediction of Shigellosis outcomes in Israel using machine learning classifiers. *Epidemiol Infect* 2018;146:1445–51. <https://doi.org/10.1017/S0950268818001498>.
- [125] Cherkasov A, et al. Use of artificial intelligence in the design of small peptide antibiotics effective against a broad spectrum of highly antibiotic-resistant superbugs. *ACS Chem Biol* 2009;4:65–74. <https://doi.org/10.1021/cb800240j>.
- [126] Ben Guebila M, Thiele I. Predicting gastrointestinal drug effects using contextualized metabolic models. *PLoS Comput Biol* 2019;15:e1007100. <https://doi.org/10.1371/journal.pcbi.1007100>.
- [127] Gao M, Igata H, Takeuchi A, Sato K, Ikegaya Y. Machine learning-based prediction of adverse drug effects: an example of seizure-inducing compounds. *J Pharmacol Sci* 2017;133:70–8. <https://doi.org/10.1016/j.jphs.2017.01.003>.
- [128] Khan DD, et al. A mechanism-based pharmacokinetic/pharmacodynamic model allows prediction of antibiotic killing from MIC values for WT and mutants. *J Antimicrob Chemother* 2015;70:3051–60. <https://doi.org/10.1093/jac/dkv233>.
- [129] Chen X, et al. Failure of internal quality control in detecting significant reagent lot shift in serum creatinine measurement. *J Clin Lab Anal* 2019;33: e22991. <https://doi.org/10.1002/jcla.22991>.
- [130] Baker MA, et al. Automated outbreak detection of hospital-associated pathogens: value to infection prevention programs. *Infect Control Hosp Epidemiol* 2020;1–6. <https://doi.org/10.1017/jice.2020.233>. Jun 10.
- [131] Brown NM, et al. Pilot evaluation of a fully automated bioinformatics system for analysis of methicillin-resistant *Staphylococcus aureus* genomes and detection of outbreaks. *J Clin Microbiol* 2019;57:e00858-19. <https://doi.org/10.1128/JCM.00858-19>.