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. Schrenzel 3, G. Greub 4

1) Clinical Bacteriology and Mycology, University Hospital Basel, Basel, Switzerland
2) Applied Microbiology Research, Department of Biomedicine, University of Basel, Basel, Switzerland
3) Laboratory of Bacteriology, University Hospitals of Geneva, Geneva, Switzerland
4) Institute of Medical Microbiology, University Hospital Lausanne, Lausanne, Switzerland

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

A R T I C L E  I N F O

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

A B S T R A C T

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 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
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-analitics 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

| Process             | Aspect                        | Example                                                                                           | References |
|---------------------|-------------------------------|---------------------------------------------------------------------------------------------------|------------|
| Pre-analytics       | Quality control              | What is the sample quality?                                                                      | [18–20]    |
|                     |                               | – Automated measurement and feedback regarding the correct filling of blood culture bottles.     |            |
|                     |                               | – Automated assessment of sample contamination including species and clinical score               |            |
| Diagnostic stewardship | Which additional diagnostic test should be ordered?                     | [26–28]                                                                                         |            |
| Analytics           | Quality control              | How reliable is the analytical performance of a test?                                               | [129]      |
|                     |                               | – Surveillance of reagent lots performance with internal and external controls and automated reporting in connection to specific used lots of time |            |
| Imaging             |                               | Are there bacteria on the microscope slide?                                                        | [30,32,33] |
|                     |                               | – Automated image acquisition with a microscope and scan for pathogen-like structures and category |            |
| Plate reading       |                               | Is there bacterial growth on the plate?                                                            | [36–38]    |
|                     |                               | – Automated image acquisition and scan for colonies and subsequent identification (telediagnostics) |            |
| Expert system       |                               | Does the detected resistance profile make sense?                                                    | [41,42]    |
|                     |                               | – Medical validation of antibiotic resistance profiles with expert database                         |            |
| Public Health       |                               | Is there a potential outbreak?                                                                    | [130,131] |
|                     |                               | – Automated screening for pathogen similarities, e.g., resistance profile or automated bioinformatics |            |
| Post-analytics      | Highlight important data     | Is there a potential bacterial phenotype?                                                          | [43,44]    |
| Sepsis treatment    |                               | What is the best treatment 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 to collect information from various sources. The education portfolio of all laboratory personnel 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. Similarly, 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, proteomics 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-
unreadable 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 is 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://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, regualting 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. Søgaard and Dr. Vladimiria 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

[1] Wilson N, MacDonald EJ, Mansoor OD, Morgan J. In bed with Siri and Google Assistant: a comparison of sexual health advice. BMJ 2017;359:j6535. https://doi.org/10.1136/bmj.j6535.

[2] 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.

[3] Baig MM, Affifi 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/s10916-019-01365-7.

[4] 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. JMR Public Health Surveill 2018;4:6. https://doi.org/10.1111/jgh.14641.

[5] Amukele T. Current state of drones in healthcare: challenges and opportunities, J Appl Lab Med 2019;4:296–8. https://doi.org/10.1373/jaml.2015031016.

[6] 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.

[7] Comellas MJ, et al. Evaluation of a new automated, fast, and fully automated glycemic pattern detection tool. Diabetes Technol Ther 2017;19:633–40. https://doi.org/10.1089/diath.2017.0180.

[8] 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.

[9] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25:44–56. https://doi.org/10.1038/s41591-019-0307-0.

[10] 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/ehw247.

[11] 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.

[12] 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.

[13] Gruson D, Helleputte T, Rousseau P, Gruson D. Data science, artificial intelligence, Blockchain, and Internet of Things. Int J Environ Res Public Health 2019;16:3847. https://doi.org/10.3390/ijerph16203847.

[14] 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/ph12090535.

[15] Coeckull CV, Cole LA, Stone DJ. Critical care, critical data. Biomed Eng Comput Biol 2019;10. https://doi.org/10.1101/17759721985656.

[16] Coeckull CV, Cole LA, Stone DJ. Critical care, critical data. Biomed Eng Comput Biol 2019;10. https://doi.org/10.1101/17759721985656.

[17] Coeckull CV, Cole LA, Stone DJ. Critical care, critical data. Biomed Eng Comput Biol 2019;10. https://doi.org/10.1101/17759721985656.

[18] Coeckull CV, Cole LA, Stone DJ. Critical care, critical data. Biomed Eng Comput Biol 2019;10. https://doi.org/10.1101/17759721985656.

[19] Coeckull CV, Cole LA, Stone DJ. Critical care, critical data. Biomed Eng Comput Biol 2019;10. https://doi.org/10.1101/17759721985656.

[20] Coeckull CV, Cole LA, Stone DJ. Critical care, critical data. Biomed Eng Comput Biol 2019;10. https://doi.org/10.1101/17759721985656.

[21] Coeckull CV, Cole LA, Stone DJ. Critical care, critical data. Biomed Eng Comput Biol 2019;10. https://doi.org/10.1101/17759721985656.
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.medint.2018.06.002.

Price 2nd WN, Cohen IG. Privacy in the age of medical big data. Nat Med 2015;21:37–43. https://doi.org/10.1038/nmed.2015.120.

Amdir et al. The COMPARE data Hubs. Database (oxford) 2019, ba136. 2019. https://doi.org/10.1093/database/baz136.

Jackson C, Gardy JL, Shladover 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.

Saxena A, et al. Ethics preparedness: facilitating ethics review during outbreaks - recommendations from an expert panel. BMC Med Ethics 2019;20:23. https://doi.org/10.1186/s12910-019-0366-x.

Caogant DA, Brynildsrud OB. Neisseria meningitidis: using genomics to understand diversity, evolution and pathogenesis. Nut Rev Microbiol 2020;18:84–96. https://doi.org/10.1093/nrmicro/2019-02682-6. 2019.

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.

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/s40504-018-0078-2.

Casey CJ, 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.

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.

Brall C, Schroder-Back P, Maeckelberghe E. Ethical aspects of digital health technologies. Clin Microbiol Infect 2020;26:1300–1304. https://doi.org/10.1016/j.cmi.2020.02.006.

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.

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.

Peiffer-Smaja 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.

Macesic N, Polubriaginof F, Tattonetti NP. Machine learning: novel bioinformatics approaches for combating antimicrobial resistance. Curr Opin Infect Dis 2017;30:511–7. https://doi.org/10.1097/QCO.0000000000000405.

Peiffer-Smaja 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.

Wiens J, Shennys 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/cxv731.

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/dky314.

Burton RJ, Albur M, Ebel 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/s12911-019-0678-9.

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.

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.

Chen ML, et al. Beyond multidrug resistance: leveraging rare variants with machine and statistical learning models in Mycobacterium tuberculosis resistance prediction. ElBioMedicine 2019;43:356–69. https://doi.org/10.1016/j.ebiomed.2019.04.016.

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.

Adak G, et al. Prediction of Shigelliosis outcomes in Israel using machine learning classifiers. Epidemiol Infect 2018;146:1445–51. https://doi.org/10.1017/S0950268818001498.

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/cb900240j.

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.

Gao M, Igata H, Takeuchi A, Sato K, Kegaya 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.

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.

Chen X, et al. Failure of internal quality control to detect agent lot shift in serum creatinine measurement. J Clin Lab Anal 2019;33:60. https://doi.org/10.1002/jcla.22991.

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.