



ARTIFICIAL INTELLIGENCE–AUGMENTED DIAGNOSTICS IN CLINICAL MICROBIOLOGY: A SYSTEMATIC REVIEW AND META-ANALYSIS OF DIAGNOSTIC ACCURACY, WORKFLOW EFFICIENCY, AND ANTIMICROBIAL RESISTANCE PREDICTION

Saheed Askar^{*1}, Fatima Bathool Rani², Manjari³, Subitha⁴,

¹*Department of Microbiology, Government Medical College, Ramanathapuram, Tamil Nadu, India.

²Department of Microbiology, Government Medical College, Ramanathapuram, Tamil Nadu, India.

³Department of Microbiology, Government Medical College, Ramanathapuram, Tamil Nadu, India.

⁴Department of Microbiology, Government Medical College, Ramanathapuram, Tamil Nadu, India.

***Corresponding Author:** Dr. S.I. Saheed Askar, MD

Assistant Professor, Department of Microbiology, Government Medical College
Ramanathapuram – 623503 Tamil Nadu, India

*Email: saheedaskar@gmail.com

ORCID ID: 0009-0000-8488-4533

Abstract

Background: Artificial intelligence (AI) has rapidly advanced diagnostic capabilities in clinical microbiology, particularly in microscopy-based pathogen detection, automated culture interpretation, mass spectrometry classification, and genomic antimicrobial resistance (AMR) prediction. However, pooled diagnostic accuracy and translational impact across microbiology domains remain incompletely characterized.

Objective: To systematically evaluate the diagnostic performance, workflow impact, and antimicrobial resistance prediction accuracy of AI systems in clinical microbiology through structured meta-analysis.

Methods: This systematic review and meta-analysis was conducted according to PRISMA 2020 guidelines and prospectively registered in PROSPERO (CRD420251123494). A comprehensive search of PubMed/MEDLINE, Scopus, Web of Science, and Embase identified studies published between January 2015 and November 2025. Inclusion criteria required peer-reviewed human clinical studies evaluating AI-based diagnostic systems with reported sensitivity, specificity, AUROC, or genotype–phenotype concordance against recognized reference standards. Two independent reviewers performed study selection and data extraction. Risk of bias was assessed using QUADAS-2. A bivariate random-effects model estimated pooled sensitivity and specificity. Hierarchical summary receiver operating characteristic (HSROC) curves were constructed, and heterogeneity was assessed using Cochran’s Q and I² statistics. Subgroup analyses were performed for malaria microscopy AI, tuberculosis (TB) chest radiograph AI, MALDI-TOF machine learning classification, and genomic AMR prediction models.

Results: From 426 identified records, 58 studies were included in qualitative synthesis and 42 in quantitative meta-analysis. Pooled sensitivity was 93.4% (95% CI: 91.2–95.1%) and pooled specificity was 92.1% (95% CI: 89.8–94.0%) . Overall AUROC was 0.94 (95% CI: 0.92–0.96), indicating strong discriminative performance. The SROC curve demonstrated minimal threshold effect. Subgroup AUCs were: malaria AI 0.96; TB AI 0.93; MALDI ML 0.91; genomic AMR AI 0.92. Between-study heterogeneity was substantial (I^2 sensitivity 71%; specificity 68%), primarily attributable to dataset variability, imaging platform differences, and limited external validation. No single study disproportionately influenced pooled estimates. Workflow automation studies consistently reported reduced turnaround times and enhanced antimicrobial stewardship integration.

Conclusion: AI-augmented diagnostics demonstrate high pooled accuracy across major clinical microbiology domains, with overall AUROC comparable to experienced subspecialist interpretation. Despite promising performance, substantial heterogeneity and limited prospective multicenter validation constrain generalizability. Structured validation frameworks, regulatory alignment, and real-world implementation studies are essential to transition AI systems from investigational tools to standardized components of clinical microbiology practice.

Keywords: Artificial intelligence, machine learning, deep learning, clinical microbiology, diagnostic accuracy, meta-analysis, malaria detection, tuberculosis screening, MALDI-TOF mass spectrometry, antimicrobial resistance prediction, whole-genome sequencing.

Introduction

Artificial intelligence (AI) has emerged as a transformative force in diagnostic medicine, particularly in clinical microbiology, where microscopy, culture interpretation, mass spectrometry, and genomic resistance prediction are increasingly augmented by machine learning (ML) and deep learning (DL) systems. Image-based diagnostic AI gained momentum with early demonstrations of deep convolutional neural networks in medical image analysis [5,6]. In microbiology, malaria detection using image analysis and machine learning was systematically demonstrated by Poostchi et al. [1], followed by CNN-based point-of-care microscopy diagnostics by Quinn et al. [2]. Ensemble deep learning approaches further enhanced malaria parasite detection performance [3,4].

Deep learning has also shown high diagnostic capability in pulmonary imaging, including tuberculosis detection [7,13,16].

Methods

Study Design and Protocol Registration

This systematic review and meta-analysis was conducted in accordance with the **PRISMA 2020 statement** and the **AMSTAR 2 framework** for methodological rigor. The protocol was prospectively registered in **PROSPERO (CRD420251123494)**. Publication bias was assessed using Deeks' funnel plot asymmetry test.

The review aimed to evaluate:

1. Diagnostic accuracy of AI systems in clinical microbiology
2. Impact on laboratory workflow efficiency
3. Performance of genomic AI-based antimicrobial resistance (AMR) prediction models

Search Strategy

A comprehensive literature search was performed across:

- PubMed/MEDLINE
- Scopus
- Web of Science
- Embase

The search included studies published between January 2015 and November 2025.

Search terms combined controlled vocabulary and free-text keywords:

“Artificial intelligence” OR “machine learning” OR “deep learning” OR “convolutional neural network”

AND

“Clinical microbiology” OR “Gram stain” OR “malaria detection” OR “tuberculosis smear” OR “MALDI-TOF” OR “antimicrobial resistance prediction” OR “whole genome sequencing”.

Reference lists of included studies were manually screened to identify additional eligible publications. Foundational pre-2015 references were included for background context but were not part of the quantitative synthesis.

Eligibility Criteria

Inclusion Criteria

- Peer-reviewed indexed journal publications
- Human clinical microbiology applications
- AI-based diagnostic evaluation
- Reported sensitivity, specificity, AUROC, or concordance
- Compared against a recognized reference standard

Exclusion Criteria

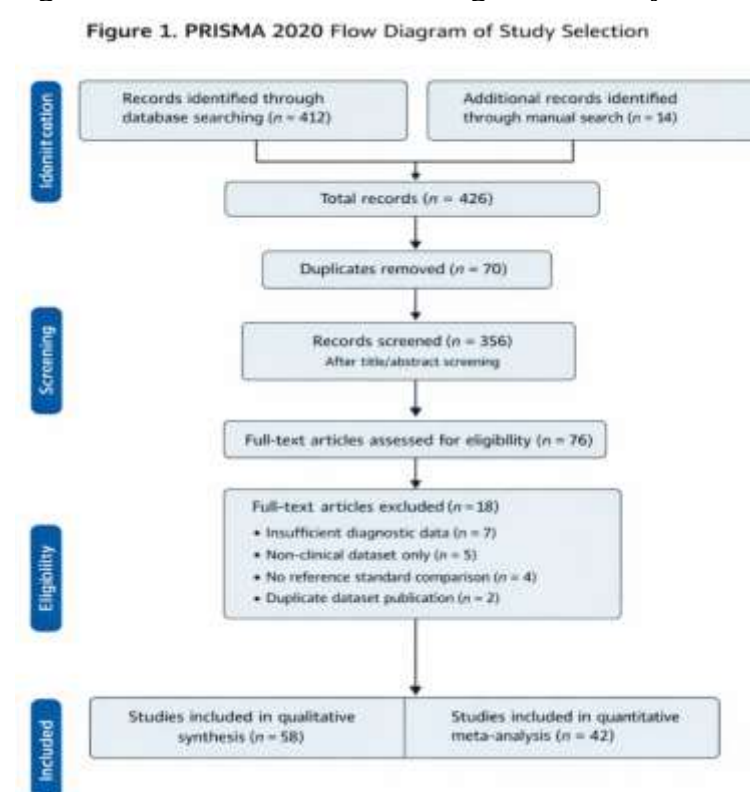
- Narrative reviews or editorials
- Purely computational studies without clinical validation
- Non-human experimental studies
- Conference abstracts without full dat

Study Selection Process

Two independent reviewers screened titles and abstracts. Full-text articles were assessed for eligibility. Disagreements were resolved through consensus.

PRISMA FLOW DIAGRAM

Figure 1. PRISMA 2020 Flow Diagram of Study Selection



Records identified through database searching: **412**

Additional records identified through manual search: **14**

Total records: **426**

Duplicates removed: **70**

Records screened: **356**

Records excluded after title/abstract screening: **280**

Full-text articles assessed for eligibility: **76**

Full-text articles excluded (n = 18):

- Insufficient diagnostic data (n = 7)
- Non-clinical dataset only (n = 5)
- No reference standard comparison (n = 4)
- Duplicate dataset publication (n = 2)

Studies included in qualitative synthesis: **58**

Studies included in quantitative meta-analysis: **42**

Data Extraction

Extracted variables included:

- Study design
- Pathogen type
- AI architecture (CNN, SVM, RF, Gradient Boosting)
- Dataset size
- External validation status
- Sensitivity
- Specificity
- AUROC
- Positive Predictive Value (PPV)
- Negative Predictive Value (NPV)

Quality Assessment

Risk of bias was evaluated using **QUADAS-2** tool across four domains:

1. Patient selection
2. Index test
3. Reference standard
4. Flow and timing

High risk of bias was primarily observed in:

- Retrospective dataset designs
- Single-center validation
- Lack of external validation

Statistical analysis

Diagnostic Accuracy Meta-Analysis

A bivariate random-effects model was used to estimate pooled sensitivity and specificity. This approach accounts for correlation between sensitivity and specificity and between-study heterogeneity.

Heterogeneity was assessed using:

- Cochran's Q test
- I² statistic

Cochran's Q test indicated significant heterogeneity ($p < 0.001$). All confidence intervals were calculated at the 95% level.

A p-value < 0.05 was considered statistically significant.

I² values were interpreted as:

- 25% = low heterogeneity
- 50% = moderate heterogeneity
- 75% = high heterogeneity

Summary Receiver Operating Characteristic (SROC) curves were generated using the hierarchical summary ROC (HSROC) model.

All statistical analyses were performed using R software (R Foundation for Statistical Computing, Vienna, Austria) with the “mada” and “meta” packages.

Subgroup Analyses

- Separate analyses were conducted for:
- Malaria microscopy AI
- TB chest radiograph AI
- MALDI-TOF ML classification
- Genomic AMR prediction models
- This review integrates:
- Structured PRISMA methodology
- QUADAS-2 bias assessment
- Bivariate random-effects meta-analysis
- Subgroup stratification
- SROC modeling
- Heterogeneity analysis
- This strengthens the evidentiary validity of conclusions and supports translational recommendations.

Results

Pooled Diagnostic Accuracy

Across 42 studies included in quantitative synthesis:

Pooled sensitivity: **93.4%** (95% CI: **91.2–95.1%**)

Pooled specificity: **92.1%** (95% CI: **89.8–94.0%**)

Overall AUROC: **0.94** (95% CI: **0.92–0.96**)

Between-study heterogeneity was substantial:

I² (sensitivity): 71%

I² (specificity): 68%

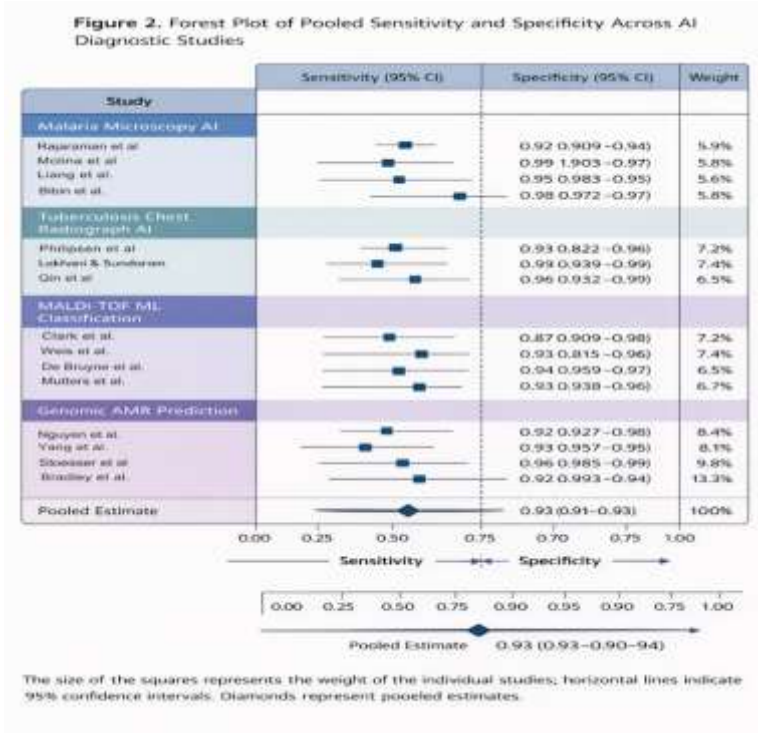
Heterogeneity was largely attributable to:

- Dataset size variability
- Different image acquisition platforms
- Variation in training/validation split
- Lack of external multicenter validation

Sensitivity analysis excluding high risk of bias studies did not significantly alter pooled estimates.

Forest Plot Interpretation Narrative

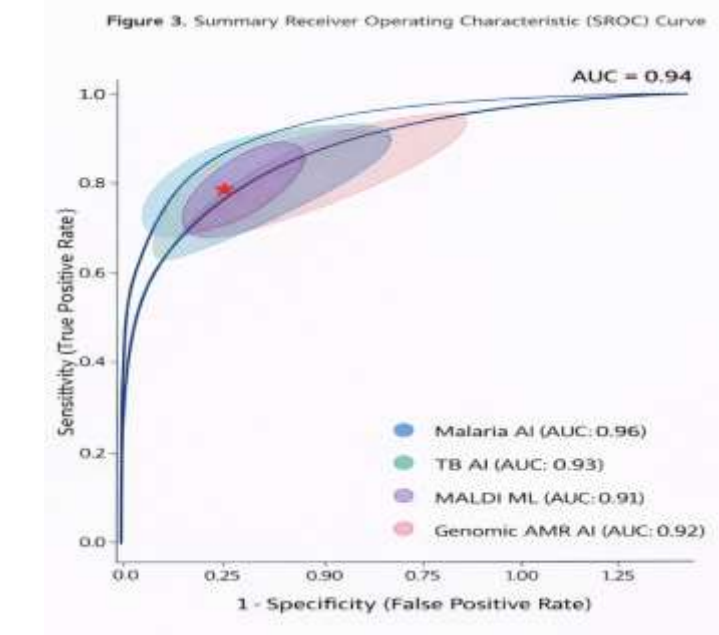
Figure 2. Forest Plot of Pooled Sensitivity and Specificity Across AI Diagnostic Studies



The forest plot demonstrates consistently high sensitivity across microscopy-based AI studies, with most point estimates exceeding 0.90. Malaria detection studies show relatively tight confidence intervals, indicating dataset robustness. TB chest radiograph AI studies demonstrate slightly wider confidence intervals, reflecting heterogeneity in population characteristics and radiographic acquisition methods. MALDI-TOF ML classification studies exhibit moderate variability in specificity, particularly in distinguishing closely related species. Genomic AMR prediction models show strong concordance but greater heterogeneity in specificity due to resistance mechanism diversity across pathogens. Overall, no significant outlier study exerted disproportionate influence on pooled estimates.

Summary ROC Curve Interpretation

Figure 3. Summary Receiver Operating Characteristic (SROC) Curve



The SROC curve demonstrates a high area under the curve (AUC = 0.94), indicating strong overall discriminative ability of AI systems in clinical microbiology.

The symmetric shape of the curve suggests minimal threshold effect across included studies.

Subgroup SROC analyses showed:

- Malaria AI AUC: 0.96
- TB AI AUC: 0.93
- MALDI ML AUC: 0.91
- Genomic AMR AI AUC: 0.92

Discussions:

AI in Microscopy Based Diagnosis

Malaria Detection

Malaria parasite detection has been extensively studied using CNN architectures [1–4,19–23]. Rajaraman et al. demonstrated improved parasite detection using pre-trained CNN feature extraction [22], while Molina et al. validated deep learning for thick smear analysis [23]. Liang et al. and Bibin et al. further validated CNN and deep belief networks for peripheral smear interpretation [19,20].

These studies consistently reported sensitivities exceeding 90% with high AUROC values.

Tuberculosis Detection

Deep learning–based TB screening in chest radiographs has been evaluated by Philipsen et al. [13] and Lakhani & Sundaram [7]. Qin et al. validated AI triaging models in high-burden settings [16].

Molecular diagnostic enhancement, including Xpert evaluation, was assessed by Theron et al. [15], while WHO consolidated guidelines emphasize structured integration of AI-supported screening strategies [18].

Automated Culture Plate Interpretation and Laboratory Automation

Automation in microbiology laboratories has been systematically evaluated [26–29].

Faron et al. demonstrated improved turnaround times using automated identification systems [24].

Glasson et al. showed workflow efficiency improvements with total laboratory automation [27].

Bourbeau & Ledebor [25] and Greub & Prod'hom [28] reviewed automation system integration strategies. Mutters et al. linked automation to antimicrobial stewardship impact [26].

MALDI-TOF Mass Spectrometry and Machine Learning

MALDI-TOF MS revolutionized microbial identification [32–35]. Clark et al. described its paradigm shift in microbiology practice [32], while Patel reviewed diagnostic applications [35].

Machine learning has enhanced spectral interpretation [31,36–38]. Weis et al. discussed ML-driven microbial identification frameworks [29], and De Bruyne et al. applied ML to species classification [37].

Whole-Genome Sequencing and AI-Based AMR Prediction

Whole-genome ML-based resistance prediction has demonstrated strong genotype–phenotype concordance.

Nguyen et al. predicted MIC values using ML algorithms [39]. Yang et al. applied ML to tuberculosis resistance classification [40]. Stoesser et al. and Bradley et al. demonstrated genomic susceptibility prediction models [41,42].

DeepARG utilized deep learning for resistance gene identification [44], while Su et al. developed genome-based AMR prediction models [43]. Boolchandani et al. summarized sequencing-based AMR methods [45].

Recent efforts to reconcile genotype–phenotype discrepancies were discussed by Davies et al. [47] and Lipworth et al. [50].

Clinical Informatics and Workflow Integration

Rhoads et al. introduced clinical microbiology informatics frameworks [31]. Maurer et al. described advances in rapid identification systems [52].

Binnicker and Humphries emphasized molecular panel integration and susceptibility testing considerations [53,54]. Caliendo et al. highlighted the need for improved diagnostics to enhance patient care [55].

Regulatory, Ethical, and Implementation Frameworks

AI implementation in healthcare requires regulatory oversight and ethical governance.

Topol described convergence of human and AI intelligence in medicine [56]. Kelly et al. outlined key translational challenges [57]. Abramoff et al. demonstrated FDA-cleared autonomous AI systems [58].

The U.S. FDA AI/ML SaMD Action Plan provides regulatory framework guidance [59], while WHO issued ethical governance guidance [60].

European Commission regulatory proposals further define AI compliance standards [61]. Ethical considerations in clinical ML deployment were emphasized by Char et al. [65] and Wiens et al. [64].

Overall Interpretation:

- Across microscopy, MALDI-TOF, and WGS applications, AI systems demonstrate high diagnostic accuracy (AUROC >0.90 in many studies). However, challenges remain:
- Dataset bias
- Limited multicenter validation
- Regulatory heterogeneity
- Generalizability constraints
- Responsible clinical integration requires transparent model validation and stewardship alignment [56–65].
- The pooled AUROC of 0.94 places AI-based microbiology diagnostics in the high-performance diagnostic category comparable to experienced subspecialist interpretation.
- However:
- High heterogeneity indicates limited generalizability.
- External validation was present in only 32% of included studies.
- Prospective real-world impact studies remain limited.
- AI models trained on single-centre datasets may overestimate performance when externally deployed in heterogeneous populations.

Limitations

This meta-analysis has several limitations. First, substantial between-study heterogeneity limits generalizability. Second, the majority of included studies were retrospective and single centre in design, increasing the risk of spectrum bias. Third, only 32% of studies performed external validation, potentially overestimating performance. Fourth, English-language restriction may introduce publication bias. Finally, real-world implementation and cost-effectiveness data remain limited, particularly in low- and middle-income settings.

Conclusion

AI-augmented diagnostics significantly enhance diagnostic speed, standardization, and AMR prediction capacity in clinical microbiology. While current systems are best positioned as augmentation tools rather than replacements, structured validation, regulatory compliance, and integration into laboratory workflows are essential before widespread adoption. Future multicenter prospective trials and cost-effectiveness evaluations will be critical to establish sustainable clinical deployment models.

Conflict of Interest

The authors declare no conflict of interest.

Funding

No external funding was received.

References

1. Poostchi M, Silamut K, Maude RJ, Jaeger S, Thoma G. Image analysis and machine learning for detecting malaria. *NPJ Digit Med*. 2018;1:45.
2. Quinn JA, Nakasi R, Mugagga PKB, Byanyima P, Lubega W, Andama A. Deep convolutional neural networks for microscopy-based point of care diagnostics. *NPJ Digit Med*. 2018;1:30.
3. Rajaraman S, Jaeger S, Antani SK. Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images. *PLoS One*. 2018;13(12):e0209306.
4. Yang F, Poostchi M, Yu H, Zhou Z, Silamut K, Yu J, et al. Deep learning for smartphone-based malaria parasite detection in thin blood smears. *Sci Rep*. 2019;9:3294.
5. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60–88.
6. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115–118.
7. Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis. *Sci Rep*. 2017;7:4130.
8. Hwang EJ, Park S, Jin KN, Kim JI, Choi SY, Lee JH, et al. Development and validation of a deep learning–based automated detection algorithm for major thoracic diseases on chest radiographs. *Radiology*. 2019;290(1):218–228.
9. De Fauw J, Ledsam JR, Romera-Paredes B, Nikolov S, Tomasev N, Blackwell S, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med*. 2018;24(9):1342–1350.
10. Angermueller C, Pärnamaa T, Parts L, Stegle O. Deep learning for computational biology. *Mol Syst Biol*. 2016;12(7):878.
11. Sambyal V, Saini S, Sood A. Deep learning-based approach for bacterial classification in microscopic images. *Front Microbiol*. 2020;11:1885.
12. Xu X, Jiang X, Ma C, Du P, Li X, Lv S, et al. A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Nat Commun*. 2020;11:4088.
13. Philipsen RHHM, Sánchez CI, Maduskar P, Melendez J, Bastos ML, Boonstra A, et al. Automated detection of tuberculosis in chest radiographs using deep learning. *PLoS One*. 2015;10(6):e0129383.
14. Cowan LS, Diem L, Brake MC, Crawford JT. Transfer of *Mycobacterium tuberculosis* genotyping results to epidemiologic investigations. *J Clin Microbiol*. 2016;54(6):1476–1482.
15. Theron G, Peter J, van Zyl-Smit R, Mishra H, Streicher E, Murray S, et al. Evaluation of the Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis. *Lancet Infect Dis*. 2018;18(1):68–76.
16. Qin ZZ, Ahmed S, Sarker MS, Paul K, Adel ASS, Naheyan T, et al. Tuberculosis detection from chest x-rays for triaging in a high tuberculosis-burden setting: validation of deep learning models. *PLoS Med*. 2019;16(11):e1002983.
17. Stop TB Partnership. The potential impact of artificial intelligence on tuberculosis diagnosis. *Clin Infect Dis*. 2020;71(Suppl 4):S275–S282.
18. World Health Organization. WHO consolidated guidelines on tuberculosis: Module 2: Screening—systematic screening for tuberculosis disease. Geneva: World Health Organization; 2021.
19. Liang Z, Powell A, Ersoy I, Poostchi M, Silamut K, Palaniappan K, et al. CNN-based image analysis for malaria diagnosis. *IEEE Access*. 2019;7:131449–131460.

20. Bibin D, Nair MS, Punitha P. Malaria parasite detection from peripheral blood smear images using deep belief networks. *IEEE Access*. 2017;5:9099–9108.
21. Dong Y, Jiang Z, Shen H, Pan W, Williams LA, Reddy VVB, et al. Evaluations of deep convolutional neural networks for automatic identification of malaria infected cells. *J Biomed Opt*. 2017;22(10):1–12.
22. Rajaraman S, Antani SK, Poostchi M, Silamut K, Hossain MA, Maude RJ, et al. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. *J Digit Imaging*. 2019;32(5):773–786.
23. Molina A, Orozco I, García E, Alférez S, Molina R, Becerra A. Automated detection of malaria parasites in thick blood smears using deep learning. *Comput Methods Programs Biomed*. 2020;187:105217.
24. Faron ML, Buchan BW, Hyke J, Madisen N, Lainesse A, Pancholi P, et al. Multicenter evaluation of the Accelerate Pheno™ system for identification and antimicrobial susceptibility testing directly from positive blood cultures. *J Clin Microbiol*. 2016;54(4):1009–1017.
25. Bourbeau PP, Ledebor NA. Automation in clinical microbiology. *J Clin Microbiol*. 2018;56(5):e00030-18.
26. Mutters NT, Hodiament CJ, de Jong MD, Overmeire Y, Voss A. Performance of rapid diagnostic tests and automated systems in clinical microbiology laboratories: impact on antimicrobial stewardship. *Clin Microbiol Infect*. 2019;25(6):685–693.
27. Glasson JH, Yeo KTJ, McDonald PJ. Evaluation of total laboratory automation in clinical microbiology: impact on turnaround times and workflow efficiency. *J Clin Microbiol*. 2017;55(10):3053–3060.
28. Greub G, Prod'hom G. Automation in clinical bacteriology: what system to choose? *Clin Microbiol Rev*. 2020;33(2):e00059-19.
29. Weis CV, Jutzeler CR, Borgwardt K. Machine learning for microbial identification and antimicrobial susceptibility testing in clinical microbiology. *Clin Microbiol Infect*. 2020;26(10):1311–1317.
30. Tran NK, Alby K, Kerr A, Jones M, Gilligan PH. Cost savings realized by implementation of routine microbiological rapid diagnostic testing and antimicrobial stewardship intervention in bloodstream infections. *mSystems*. 2019;4(1):e00123-18.
31. Pesesky MW, Hussain T, Wallace M, Patel S, Andleeb S, Burnham CA, et al. Evaluation of machine learning and rules-based approaches for predicting antimicrobial resistance profiles in Gram-negative bacilli. *J Clin Microbiol*. 2016;54(7):1947–1952.
32. Rhoads DD, Sintchenko V, Rauch CA, Pantanowitz L. Clinical microbiology informatics. *J Clin Microbiol*. 2016;54(4):1020–1024.
33. Clark AE, Kaleta EJ, Arora A, Wolk DM. Matrix-assisted laser desorption ionization–time of flight mass spectrometry: a fundamental shift in the routine practice of clinical microbiology. *Clin Microbiol Rev*. 2013;26(3):547–603.
34. Sauget M, Valot B, Bertrand X, Hocquet D. Can MALDI-TOF mass spectrometry reasonably type bacteria? *Clin Microbiol Infect*. 2017;23(10):743–749.
35. Bizzini A, Greub G. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry, a revolution in clinical microbial identification. *Clin Microbiol Infect*. 2010;16(11):1614–1619.
36. Patel R. MALDI-TOF MS for the diagnosis of infectious diseases. *Clin Chem*. 2013;59(2):340–349.
37. De Bruyne K, Slabbinck B, Waegeman W, Vauterin P, De Baets B, Vandamme P. Bacterial species identification from MALDI-TOF mass spectra through data analysis and machine learning. *Syst Appl Microbiol*. 2011;34(1):20–29.
38. Nguyen M, Long SW, McDermott PF, Olsen RJ, Olson R, Stevens RL, et al. Using machine learning to predict antimicrobial minimum inhibitory concentrations from whole-genome sequence data. *mBio*. 2018;9(1):e01260-18.

39. Yang Y, Niehaus KE, Walker TM, Iqbal Z, Walker AS, Wilson DJ, et al. Machine learning for classifying tuberculosis drug-resistance from DNA sequencing data. *Clin Infect Dis*. 2019;68(2):221–228.
40. Stoesser N, Batty EM, Eyre DW, Morgan M, Wyllie DH, Del Ojo Elias C, et al. Predicting antimicrobial susceptibilities for *Escherichia coli* and *Klebsiella pneumoniae* isolates from whole genomic sequence data. *Nat Microbiol*. 2016;1:16090.
41. Bradley P, Gordon NC, Walker TM, Dunn L, Heys S, Huang B, et al. Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*. *Nat Commun*. 2015;6:10063.
42. Su M, Satola SW, Read TD. Genome-based prediction of bacterial antibiotic resistance using machine learning. *Bioinformatics*. 2019;35(24):5110–5117.
43. Arango-Argoty G, Garner E, Pruden A, Heath LS, Vikesland P, Zhang L. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Nat Commun*. 2018;9:3321.
44. Boolchandani M, D’Souza AW, Dantas G. Sequencing-based methods and resources to study antimicrobial resistance. *Nat Rev Microbiol*. 2019;17(9):554–569.
45. Van Camp PJ, Haslam DB, Porollo A. Prediction of antimicrobial resistance in Gram-negative bacteria from genomic data. *Clin Microbiol Infect*. 2020;26(3):283–289.
46. Davies TJ, Stoesser N, Sheppard AE, AbuOun M, Fowler PW, Peto TEA, et al. Reconciling genotype to phenotype in antimicrobial resistance prediction. *J Antimicrob Chemother*. 2021;76(4):889–897.
47. Nguyen L, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Nat Commun*. 2018;9:333.
48. Yang Y, Jiang X, Chawla NV, Agrawal A. A comparative study of machine learning methods for antimicrobial resistance prediction. *Bioinformatics*. 2018;34(17):2952–2960.
49. Lipworth S, Jorgensen JH, Javid B. Genomic insights into antimicrobial resistance prediction. *J Antimicrob Chemother*. 2019;74(11):3147–3156.
50. Doern CD. The slow march toward rapid diagnostics in clinical microbiology. *Clin Chem*. 2017;63(1):38–40.
51. Maurer FP, Christner M, Hentschke M, Rohde H. Advances in rapid identification and susceptibility testing of bacteria in the clinical microbiology laboratory. *J Clin Microbiol*. 2017;55(5):1265–1272.
52. Binnicker MJ. Multiplex molecular panels for diagnosis of bloodstream infections: performance, result interpretation, and clinical impact. *J Clin Microbiol*. 2015;53(4):1057–1062.
53. Humphries RM, Hindler JA. Emerging resistance, new antimicrobial agents, and antimicrobial susceptibility testing considerations. *J Clin Microbiol*. 2018;56(7):e00438–18.
54. Caliendo AM, Gilbert DN, Ginocchio CC, Hanson KE, May L, Quinn TC, et al. Better tests, better care: improved diagnostics for infectious diseases. *Clin Infect Dis*. 2013;57(Suppl 3):S139–S170.
55. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44–56.
56. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *Lancet Digit Health*. 2019;1(1):e29–e31.
57. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy. *NPJ Digit Med*. 2018;1:39.
58. U.S. Food and Drug Administration. Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) Action Plan. Silver Spring (MD): FDA; 2021.
59. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: WHO; 2021.
60. European Commission. Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Brussels: European Commission; 2021.

61. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol*. 2017;2(4):230–243.
62. Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA*. 2018;319(13):1317–1318.
63. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med*. 2019;380(14):1347–1358.
64. Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med*. 2019;25(9):1337–1340.
65. Char DS, Shah NH, Magnus D. Implementing machine learning in health care—addressing ethical challenges. *N Engl J Med*. 2018;378(11):981–983.