Biomedical Instrumentation 2 (2026) 100052

Contents lists available at ScienceDirect

Biomedical Instrumentation

journal homepage: www.keaipublishing.com/en/journals/biomedical-instrumentation/

Review article

Artificial intelligence (AI) for polymerase chain reaction (PCR): A state-of-the-art review

Stephanie Andaluz[a,c], Aman Lv[a,b], Pengyu Yu[a], Wenhao Hui[d], Ren Shen[a,b], Tzu-Ming Liu[c], Miguel Brito[e], Pui-In Mak[a,b], Rui P. Martins[a,b], Yanwei Jia[a,b,*]

[a] *State Key Laboratory of Analog and Mixed-Signal VLSI, Institute of Microelectronics, University of Macau, Macao Special Administrative Region of China*
[b] *Faculty of Science and Technology, University of Macau, Macao Special Administrative Region of China*
[c] *Faculty of Health Sciences, University of Macau, Macao Special Administrative Region of China*
[d] *Zhuhai ProMed Precision Medical Technology Co., Ltd, Zhuhai 519031, China*
[e] *H&TRC – Health & Technology Research Center, ESSL – Escola Superior de Saúde de Lisboa, Polytechnic University of Lisbon, Lisbon 1990-096, Portugal*

A R T I C L E I N F O

Keywords:
Artificial intelligence (AI)
Machine learning (ML)
Polymerase chain reaction (PCR)
DNA detection

A B S T R A C T

Polymerase chain reaction (PCR) and its advanced derivatives—quantitative PCR (qPCR), digital PCR (dPCR), high-resolution melting (HRM) analysis, and isothermal amplification—remain central to molecular diagnostics. Their growing data complexity demands computational solutions beyond traditional analysis. Meanwhile, the advancement of artificial intelligence (AI) algorithms has driven progress from conventional shallow machine learning (ML) to more complex deep learning approaches. Currently, AI is able to provide powerful frameworks for interpreting amplification dynamics, optimizing assay design, and visualizing molecular reactions in real time. With the assistance of AI, PCR can be transformed from a common laboratory technique into an intelligent diagnostic system. Herein, we review recent progress at the intersection of AI and PCR across biomedical and clinical domains. These studies demonstrate that AI-enhanced PCR platforms have significantly improved diagnostic accuracy, reproducibility, and analytical throughput, while simultaneously reducing operator dependency and cost. AI-enabled PCR is poised to become a cornerstone for next-generation, intelligent molecular diagnostics in medicine.

1. Introduction

Polymerase chain reaction (PCR) is a gold-standard technique in molecular biology for DNA detection and analysis. Its applications span countless fields, including biomedical research, clinical diagnostics, agricultural biotechnology, environmental surveillance, and forensic science [1]. A typical PCR reaction mixture contains template DNA, DNA polymerase, forward and reverse primers, deoxynucleotide triphosphates (dNTPs), and a magnesium-containing buffer. The forward and reverse primers are designed to bind to a specific region of the template DNA. At a high melting temperature (Tm) (94 °C – 98 °C), the double-stranded DNA (dsDNA) is denatured into two single-stranded DNAs (ssDNA). When the temperature is quickly lowered to an annealing temperature (55 °C – 65 °C), each primer is annealed to its specific ssDNA target and extended by adding dNTPs in the presence of the polymerase. Then the bound primer is extended to the full length of the template DNA at the extension temperature (72 °C). With one thermal cycle, one copy of the dsDNA target is doubled into two copies of dsDNA. Continuing thermal cycling can amplify the target DNA copy number exponentially, accumulating to billions of copies of DNA in 50 cycles. With fluorescence dye or probe, the amplification can be monitored in real time or at the end point, rendering PCR a sensitive, robust, and specific method for DNA identification.

However, the successful implementation of PCR requires careful and methodologically rigorous design, including the optimization of numerous interdependent parameters, primer selection, target region specificity, Tms, GC content, secondary structure avoidance, and reaction thermodynamics [2]. These design considerations become increasingly complex, especially in quantitative PCR (qPCR), multiplex PCR, and digital PCR (dPCR), where reaction dynamics and molecular interactions must be precisely controlled to ensure analytical accuracy [3]. Besides, the analysis of amplification data introduces an additional

Peer review under responsibility of Xi'an Jiaotong University.

* Corresponding author at: State Key Laboratory of Analog and Mixed-Signal VLSI, Institute of Microelectronics, University of Macau, Macao Special Administrative Region of China.

E-mail address: yanweijia@um.edu.mo (Y. Jia).

https://doi.org/10.1016/j.bmi.2026.100052
Received 11 May 2026; Received in revised form 10 June 2026; Accepted 21 June 2026
Available online 2 July 2026

3050-5372/© 2026 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

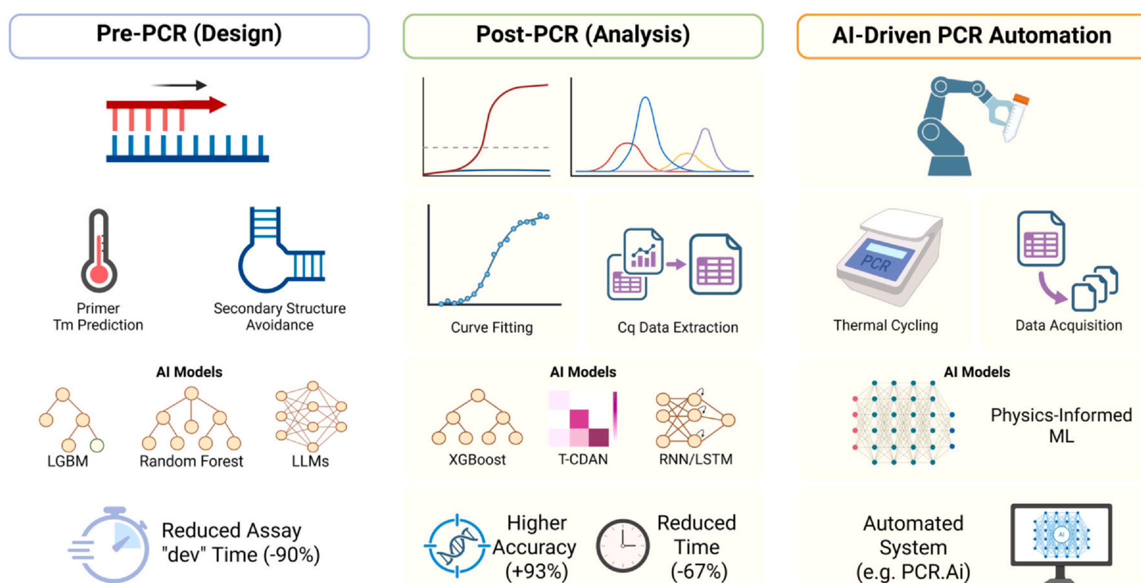


Fig. 1. AI-enhanced PCR workflow. Application of AI models at the three key stages of PCR. AI models applied before PCR (LGBM, Random Forest, LLMs) improve primer design and reduce assay development time up to 90%. After PCR (XGBoost, T-CDAN, RNN/LSTM), analysis accuracy increases up to 93%. Fully automated systems (e.g., PCR.Ai) integrate thermal cycling and data acquisition with physics-informed ML, enhancing both the efficiency and accuracy of PCR workflows. Note: This figure was created with BioRender.

layer of complexity. Interpreting amplification kinetics, quantifying nucleic acid abundance, and distinguishing true signal from background noise require robust analytical frameworks and a solid understanding of fluorescence chemistry, reaction efficiencies, and statistical variability.

In handling complex reaction systems and data analysis, artificial intelligence (AI) and machine learning (ML) have demonstrated their strong power. Modern AI systems are defined by several different core technical specifications, including advanced pattern recognition, autonomous reasoning, and adaptive learning capabilities [4–8]. On a broad scale, these architectures can be subdivided into classes according to their functional capabilities, ranging from reactive models that provide only output in response to specific inputs to limited memory models that are capable of learning from historical data, as well as novel agentic systems designed toward goal-oriented planning and multi-step execution with minimal human intervention [6,9,10]. Additionally, the evolution of multimodal large language models (MLLMs) and reasoning-based architectures has enabled AI to process heterogeneous data sources while performing complex logical inference through chain-of-thought reasoning [11].

ML algorithms have demonstrated remarkable capabilities in processing large datasets to identify patterns that would be difficult for humans to discern. In the context of PCR, these algorithms can predict optimal reaction conditions, design specific primers and probes, and interpret complex amplification curves with unprecedented accuracy. The application of deep learning techniques, including convolutional neural networks and transformer-based models, has opened new possibilities for multiplex PCR analysis and real-time pathogen detection, particularly during global health emergencies like the COVID−19 pandemic [12,13].

The deployment of these technical capabilities is governed by foundational ethical principles intended to ensure responsible and trustworthy AI implementation [14]. These principles mandate transparency and explainability, ensuring that AI-driven decisions are interpretable to users; fairness and non-discrimination, requiring the active mitigation of algorithmic bias; and robust privacy protection and data security throughout the AI lifecycle [14,15].

So far, AI has been widely used for pattern recognition and complex phenomena modeling, an advance that has opened new frontiers in genomics, structural biology, pharmacology, and systems biology [16]. Given the robustness and reproducibility of PCR process, a rational AI algorithm can be developed to enhance accuracy in primer design, tune the thermal cycling protocol, improve the data analysis strategies, and reduce the operator-dependent variabilities. In this paper, we review the current AI models that have been developed for PCR in pre-PCR design, post-PCR analysis, and the automation of results interpretation (Fig. 1).

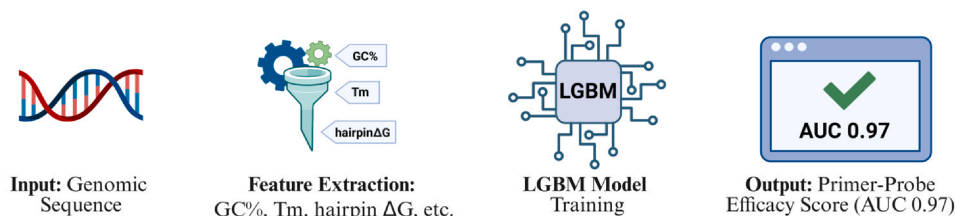
The AI models successfully applied for PCR fall into three functional families. Tree-based models (LGBM, Random Forest, XGBoost) excel at tabular feature data (e.g., primer GC%, Tm, hairpin ΔG) and are highly interpretable via feature importance. Recurrent neural networks (RNN, LSTM, GRU) are designed for time-series amplification curves, enabling early prediction of positivity. Domain adaptation and physics-informed approaches (T-CDAN, physics-informed ML) improve cross-platform generalizability but require more complex training pipelines. These families are summarized in Table 1 and will be discussed throughout the review. This comprehensive literature review aims to provide a curated library of AI tools for PCR, with the goal of laying the foundation for an integrated intelligent system for molecular diagnostics.

2. AI-Driven Pre-PCR Optimization

Three interconnected pain points determine PCR success or failure even before the first thermal cycle: First, primer design remains a complex multi-parameter optimization problem, balancing GC content, Tm, secondary structure, and self-complementarity. These variables are interrelated, and simple heuristic rules are insufficient for complex or AT-rich genomes. Second, mispriming-induced off-target amplification can lead to false positives or failed experimental runs. Still, typical BLAST-based specificity checking is generally too coarse to capture all the cross-reactivity observed in practice. Finally, manual assay development typically takes weeks to months, an impossible timescale during an outbreak. AI solves these problems by learning patterns from large databases of primers, predicting binding specificity with accuracy beyond that of thermodynamic models alone, and automating parameter optimization, often achieving development time reductions of up to 90% [13].

Table 1
Overview of AI models used in PCR studies.

AI Model	Core Principle (plain English)	Typical PCR Application	Key Advantage	Limitation (in PCR context)
LGBM (Light Gradient Boosting Machine)	Fast decision-tree boosting that learns from errors sequentially; handles missing data well	Primer-probe efficacy prediction (BioInnovate AI)	Very fast training; works on large datasets; AUC up to 0.97	Black-box; less interpretable than simple logistic regression
Random Forest	Builds many decision trees and averages their votes to reduce overfitting	Primer parameter optimization (AutoPVPrimer)	Robust to outliers; outputs which features (e.g., GC%, Tm) matter most	Can overfit on small primer datasets (< 500 pairs)
XGBoost	Sequential tree boosting with regularization to prevent overfitting	Duplex qPCR classification (VC-PCR system)	State-of-the-art accuracy on tabular fluorescence features	Requires careful hyperparameter tuning; slower than LGBM on very large data
T-CDAN (Transformer-based Conditional Domain Adversarial Network)	Learns instrument-invariant features by confusing a domain classifier	Cross-platform amplification curve classification	Generalizes across PCR machines and reagent lots without retraining	Computationally heavy; needs paired data from source and target domains
RNN (Recurrent Neural Network)	Processes sequences cycle-by-cycle, maintaining a “memory” of past cycles	Early prediction of PCR positivity (13 cycles vs. 40)	Captures temporal amplification patterns; reduces run time by 67.5%	Requires time-series data; can be overkill for end-point PCR
LSTM (Long Short-Term Memory)	A type of RNN that avoids “forgetting” long-term dependencies	Same as RNN (often used together)	Better at learning very long cycle sequences (> 50 cycles)	More parameters → needs more training data
GRU (Gated Recurrent Unit)	A simpler, faster variant of LSTM	Same as RNN/LSTM	Computationally lighter than LSTM; similar performance	Slightly less expressive than LSTM for very complex patterns
SVM (Support Vector Machine)	Finds the optimal boundary separating two classes (e.g., positive vs. negative)	COVID−19 breath sample classification (with feature extraction)	Works well on small datasets; interpretable (linear kernel)	Struggles with very nonlinear curve shapes; requires feature engineering
ANN (Artificial Neural Network)	A network of simple computing units (“neurons”) that learn nonlinear patterns	Tm prediction for thermostable proteins (custom ANN)	Can model complex, nonlinear relationships	Black-box; needs large datasets to avoid overfitting
MLP (Multilayer Perceptron)	A basic type of ANN with one or more hidden layers	Same as ANN (often used interchangeably)	Simple to implement; good for tabular data	Less powerful than deep learning for image/time-series data
Domain Adaptation (general concept, not a single model)	Adjusts a model trained on one distribution (e.g., synthetic DNA) to work on another (e.g., clinical isolates)	T-CDAN (above) and other cross-platform methods	Enables model reuse across labs/ instruments	Requires access to some target-domain labelled data
Physics-informed ML (hybrid approach)	Embeds known physical laws (e.g., thermodynamics, reaction kinetics) into the model loss function	Absolute quantification without standard curves	More generalizable; needs less training data	Harder to implement; domain knowledge required



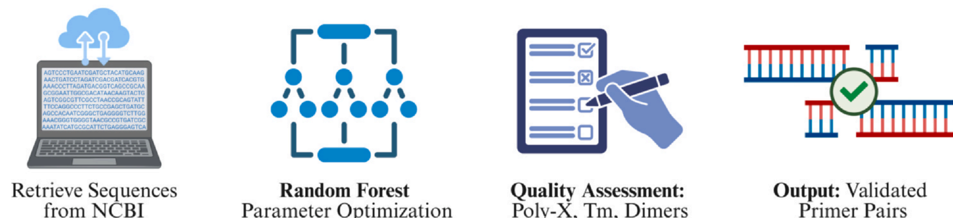


Fig. 2. AI-driven primer design workflows. (A) BioInnovate AI platform: genomic sequences are featurised (GC%, Tm, secondary structure), then a Light Gradient Boosting Machine (LGBM) classifier predicts primer-probe efficacy, achieving an AUC of 0.97 [13]. (B) AutoPVPrimer pipeline: sequences are retrieved from NCBI, optimised via a random forest classifier, quality-checked for poly-X, Tm, and primer-dimers, and output as validated primer pairs [18]. Note: This figure was created with BioRender.

2.1. Primer design: Principles and computational approaches

Primer design is the single most important step for successful PCR amplification. However, heuristic rules for primer design have, for many years, been followed with manual adjustments. The integration of various primer design parameters within one decision framework has only recently become possible with the emergence of AI-driven software platforms.

2.1.1. Intelligent primer and probe design

An example of such a system is the BioInnovate AI platform, which uses Light Gradient Boosting Machine (LGBM) models. LGBM is an open-source, very fast implementation of gradient boosting on decision trees, enabling efficient learning on large datasets (Fig. 2A). This model set reaches the highest performance level with an AUC (metric for binary classification problems in ML) of 0.97 for predicting effective primer-probe combinations [13]. The system reduces the development cycle of each PCR assay by almost 90%, enabling reagent design and optimization for a wide range of pathogens at unprecedented speed. Recently, PyPCRtool has been released, which provides a novel application of *in silico* PCR simulation and primer verification in Python, allowing users to run programs on local servers, offering unique deployment advantages for researchers [17].

AutoPVPrimer represents another example that demonstrates how AI can enhance primer design in an automated fashion, especially in the context of plant virus detection. The system employs AI and ML to automate two previously manual steps: (1) the retrieval of genomic sequences from NCBI, and (2) optimization of primer parameters (Fig. 2B). For the latter, a random forest classifier is used to guide optimization [18]. A random forest classifier builds many decision trees and combines their outputs by voting, thereby reducing the variance of individual trees and increasing accuracy. This pipeline also performs the routine quality assessment of poly-X and Tm values, but the interesting new feature is that it provides visual inspection of primer dimers.

2.1.2. Automated primer design tools

Several bioinformatic tools have been introduced for primer design. Primer3, incorporating thermodynamic calculations and salt correction formulas, represents a step forward compared to older methods [19]. In addition, recent developments in Tm calculations consider divalent ions in amounts typically present in PCR buffers; this helps minimize surprises when running experiments. Recently, ColabPCR demonstrated the cloud primer design concept using Google Collaboratory; thus, there is no need for locally installed high-performance computing hardware [20]. After the Primer3 program selects appropriate primer length and Tm, ColabPCR performs a BLASTn search to identify potential off-target binding sites. Because ColabPCR flags off-target matches before ordering primers, the primers are more likely to amplify only the intended target.

Among the versatile tools for primer design for qPCR, QuantPrime stands out; it has been employed in genomic studies requiring high-throughput quantification [21]. QuantPrime is an excellent choice when primer sets are required for a very high number of genes (e.g., transcriptomics). Moreover, QuantPrime supports the design of primers for hydrolysis probes (TaqMan) and oligonucleotide *in situ* hybridization, besides SYBR Green-based assays.

Another recent program is FastCloneAssist, which simplifies primer design for the FastCloning process, requires users to calculate overlapping PCR primers for assembling DNA fragments seamlessly [22]. By eliminating the need to perform calculations manually, the tool brings FastCloning within reach of a researcher of any experience level.

2.1.3. Machine learning-enhanced primer selection

Recent applications of ML techniques to primer optimization have shown good results. ARMSprimer3, a Python program designed for amplification refractory mutation system (ARMS)-PCR, automates the design of primers with allele-specific mismatches to enhance specificity for single-nucleotide polymorphism (SNP) detection [23]. In other words, it addresses one of the most difficult aspects of primer design, namely the design of deliberate mismatches that destabilize non-specific binding yet maintain efficient amplification of the target.

While tools like CREPE (CREate Primers and Evaluate) integrate Primer3 functionality with *in silico* PCR analysis to perform large-scale primer design and specificity assessment [24], this tool ranks primer pairs and estimates off-target binding risk. If a pipeline classifies a primer pair as "acceptable", the experimental validation yields successful amplification for more than 90% of such primer pairs, which strongly supports the automated evaluation step's validity. The output contains the best primer pairs for each target site and a measure of off-target binding potential. This high validation rate strongly supports the

idea that ML-assisted screening can replace substantial trial-and-error optimization work.

2.1.4. Tm and advanced Tm prediction models

An accurate prediction of the Tm is an important criterion in selecting primers. Indeed, the Tm is used as a guide for the specificity and efficiency of primer binding during PCR [25]. The nearest-neighbor model accounts for the interactions between neighboring base pairs. Tm calculations derived from this model incorporate enthalpy (ΔH) and entropy (ΔS) variations, as well as GC content and DNA sequence length.

HRM curve analysis of PCR products obtained from five seawater diatom species allowed the empirical derivation of formulas for predicting Tm values with an average error of less than 1°C [25]. It was demonstrated that predicted Tm values could be used for the identification of unknown samples, thus opening new avenues to address practical problems in PCR-based methods. The standard Tm models describe the thermodynamics of natural DNA duplexes. Deep learning has also been applied to the Tm prediction of modified nucleotides. Recently, Tm prediction models were developed for gapmer antisense oligonucleotides modified with amido-bridged nucleic acids (AmNAs) using a library of 157 oligonucleotides [26]. These artificial nucleic acid modifications that enhance the Tm were elucidated through differential scanning calorimetry and molecular dynamics simulations, which are useful in designing antisense oligonucleotides with an additional requirement for high thermal stability.

ML has also been successfully applied to protein design, specifically for achieving a target Tm [27]. On a nonredundant set of proteins and embeddings from state-of-the-art large protein language models (ProtBert, ProtGPT2, and ProtT5), the authors achieved a correlation of 0.89 between predicted and measured Tm. Their custom artificial neural network (ANN) and MLP regressor models form a unified predictive framework (Fig. 3). In practical terms, they can now design heat-stable enzymes with confidence, without exhaustive trial-and-error screening.

Gradient-boosting machines (LGBM, XGBoost) and random forest classifiers are the two most common AI methods employed in the five tools discussed: BioInnovate AI, AutoPVPrimer, ColabPCR, QuantPrime, and FastCloneAssist. In this respect, all studies demonstrate how the use of AI greatly reduces manual iterations in the design process. However, a common limitation across these tools lies in their reliance on well-curated training datasets that include both validated and failed primer sets, information that is rarely publicly available.

2.2. Secondary structure prediction and analysis

Understanding the secondary structure of primers, probes, and target amplicons is essential for predicting PCR efficiency and reducing non-specific amplification. Recently, there has been a major breakthrough in the deep learning-based prediction of RNA secondary structures, and the prediction accuracies of these new deep learning methods are significantly higher than those of conventional thermodynamic-based methods.

2.2.1. Hairpin and self-complementarity assessment

The formation of secondary structures in primers, probes, and amplicons plays a critical role in determining the efficiency and specificity of PCR [28]. In this study, the best primers for glutathione reductase (GR) gene expression analysis in rice were those without secondary structure in the form of hairpins and self-dimers. This was achieved by selecting several candidates and analyzing them with bioinformatic tools, including Primer3, Geneious Primer, and Primer-BLAST. A primer pair exhibiting no secondary structure and an optimal annealing temperature was identified from the initial candidates.

Other software tools that can be used for the quantitative analysis of secondary structures are OligoAnalyzer (https://www.idtdna.com) and mfold (https://www.unafold.org/mfold.php). The Tm of self-dimer or hairpin structures should be below the defined threshold temperatures. For primers, this usually indicates a weak structure, with ΔG values of

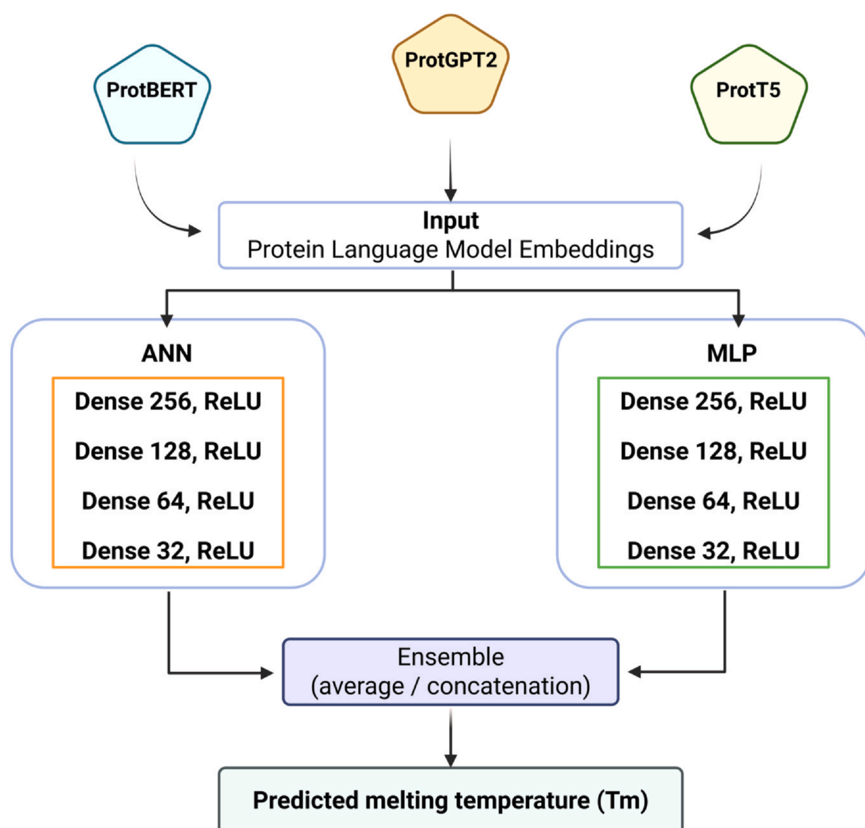


Fig. 3. Architecture of the integrated ANN and MLP framework for Tm prediction. The framework takes protein sequence embeddings from three large language models (ProtBERT, ProtGPT2, ProtT5) as input. The ANN comprises four dense hidden layers with 256, 128, 64, and 32 units, each activated by ReLU. The multilayer perception (MLP) regressor shares an identical architecture. Predictions from both models are combined via ensemble averaging to output the final predicted Tm. The ensemble approach improves correlation with experimental Tm to 0.89. Note: This figure was created with BioRender. Reprinted with permission from ref [27].

more than −9 kcal/mol considered acceptable.

2.2.2. RNA secondary structure and functional importance

Knowledge of the secondary structure of RNAs is relevant to many applications, such as guiding antisense oligonucleotide design and identifying regulatory motifs in viral genomes. SHAPE-MaP (selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling) allowed the first global structural characterization of 476 human primary microRNAs [29]. The results reveal that the structures of microRNAs differ considerably from purely in silico predictions, especially in apical loops and basal segments. This again highlights the need for experimentally validating predicted structures.

In this context, for example, in circular RNA (circRNA) applications, the secondary structures formed by the backsplice junctions (BSJ) are important regulatory elements [30]. In such cases, the self-cleavage of circRNAs induced by strategically engineered oligonucleotides was revealed by combined chemical probing and in silico modelling approaches to disrupt the secondary structures formed around the BSJs, thus providing useful insight into structure-function relationships.

2.2.3. Amplicon secondary structure effects

The folding characteristics of PCR amplicons strongly influence downstream applications. A new technology, termed “Snake” technology, uses PCR amplicons that have been structurally engineered to fold into stem-loop secondary structures to improve the cleavage of FRET probes and enhance signal production [31]. This is done by having one of the PCR primers contain a special 5'-flap sequence that causes amplicons to fold into such structures. Under these conditions, the FRET dual-labeled probes bind to the stem-loop structures and form the best possible substrates for the 5'-nuclease activity of Taq DNA polymerase. Indeed, such folding increases the cleavage of FRET probes by 5'-nuclease activity relative to cleavage under standard conditions, such as in the TaqMan assay, where cleavage occurs by different mechanisms. Increased cleavage efficiency results in greater fluorescence signal generation during PCR without affecting PCR yield. Another advantage of the Snake system is that it allows the use of shorter FRET probes with inherently lower background fluorescence, which significantly increases the sensitivity and specificity of detecting small genetic variations, such as single-nucleotide polymorphisms.

While tools such as mfold and OligoAnalyzer are standard, modern deep learning methods have surpassed traditional methods in RNA secondary structure prediction. However, this conclusion does not hold for sequences with a high divergence from the training dataset [32]. For PCR amplicon design, a pragmatic consensus from the literature is that experimental validation (e.g., SHAPE-MaP) remains necessary for critical applications, since many in silico predictions of hairpins and self-dimers do not always correspond to those observed in the wet lab, especially in GC-rich regions.

2.3. Mispriming prediction and off-target amplification prevention

Mispriming and off-target amplification are among the most significant problems in assay development for PCR and multiplex PCR in particular. With the help of AI, these problems may be predicted, considering sequence homology, thermodynamic stability, and the effect of primer concentration.

2.3.1. In silico prediction of off-target amplification

The prediction and identification of possible off-target binding sites are some of the crucial pre-PCR optimization steps [33]. A single detailed *in silico* study on problematic primers relates to the development of an *Anopheles dirus* complex species identification PCR assay (DiCSIP). The primers were analyzed in Benchling and Primer-BLAST to identify those with very high GC content and several off-target binding sites; these primers were then redesigned to enhance specificity without losing amplification efficiency for all five member species.

An extensive *in silico* genomic analysis of all published PCR primers targeting the prfA-virulence gene cluster in *Listeria* species revealed strong (> 94%) primer mapping only for the hlyA gene, whereas all other genes had very weak (< 50%) mapping results [34]. Potential reasons for nonbinding and false-negative results are nucleotide polymorphisms at the 3' end of primers, again highlighting the necessity for primer design based on the analysis of several isolates.

2.3.2. Advanced specificity verification techniques

At the genome-wide level, the CRISPR-based methods are considered the most advanced approaches for evaluating off-target activities. The iGUIDE method improves upon GUIDE-seq by filtering mispriming events to distinguish true signals, allowing specification of Cas9-guided cleavage locations [35]. Importantly, this study showed that endogenous background DNA double-strand breaks tend to occur in open chromatin, gene-rich areas, and fragile sites on the chromosome.

Another novel strategy that removes primer-dimers and suppresses misamplification from homologous sequences is RNase H-dependent PCR (rhPCR) [36]. Briefly, this method uses blocked primers containing a single ribonucleotide residue that becomes active only upon cleavage by *Pyrococcus abyssi* RNase H2, thus eliminating primer-dimer formation and substantially improving assay specificity. Moreover, mismatches around the RNA cleavage site reduce the cleavage efficiency, which further increases the specificity of the assay.

The methods thus far reviewed, spanning from in silico BLAST-based screening to rhPCR and CRISPR-based iGUIDE, share a common conclusion: specificity, (i.e., zero false positives), can only be attained with orthogonal approaches. AI is not yet able to guarantee specificity alone because primer-genome interactions occur in the context of 3D chromatin accessibility and sample matrix effects that are rarely encoded in training data. The current best practice is therefore to apply ML-assisted primer screening, followed by at least one experimental specificity validation, such as rhPCR or GUIDE-seq.

2.4. Evaluating large language models for PCR primer design

An intriguing question arises as to whether large language models (LLMs) such as GPT−3.5 and GPT−4 can be harnessed for primer design [37]. An evaluation of primers generated by LLMs for monkeypox virus detection showed that LLMs can reasonably generate sequences and predict Tms but poorly predict GC content. Of the total number of primer pairs generated, only three from GPT−4 and two from GPT−3.5 could, in theory, produce PCR products, and only one pair has suitable parameters for experimental validation. Thus, LLMs could be used for primer design, but human input is still required to ensure the quality of the assay.

LLMs (GPT−3.5, GPT−4) can generate plausible primer sequences and predict Tm with reasonable accuracy, but they fail consistently on GC content and often produce primer pairs that cannot generate PCR products. This illustrates the broad pain point that generative AI is not yet ready for production primer design unless it is fine-tuned for the task. Thus, human oversight remains essential, and hybrid LLM-based systems augmented with thermodynamic filtering hold promise for the near term.

3. AI-driven post-PCR analysis

Of the various stages of qPCR, the post-PCR data analysis stage is considered to be among the most critical and most challenging. Recent advances in AI have revolutionized post-PCR analysis, refining the extraction, analysis, and interpretation of fluorescence data generated during amplification. [38].

Three key issues make data analysis after running the PCR challenging. First, the methods to analyze amplification curves are highly variable. Fixed-threshold and first-derivative methods fail if baseline

fluorescence drifts or if there are inter-well differences in amplification efficiency. Second, the quantification cycle (Cq) value's determination depends on the instrument and the assay chemistry, making inter-laboratory comparisons problematic and even more so in multiplex reactions, where signals overlap. Third, defining positive/negative calls based on the binary presence or absence of a Cq cutoff results in false positives near the limit of detection and cannot account for run-to-run variability. An AI approach to these issues involves learning the shape of the curve, estimating the initial copy number of the target without the need for standard curves, and assigning probabilistic classes, thereby reducing operator bias.

3.1. Amplification curves analysis and interpretation

Traditional methods, such as fixed-threshold setting and first-derivative maximum (FDM) methods, are not adequate for datasets that contain variability at a high scale. The AI-assisted multiplex SYBR green qPCR system with velocity-controlled PCR integration demonstrated remarkable performance using ML frameworks combining extreme gradient boosting (XGBoost). XGBoost is a powerful algorithm that builds a predictive model by combining many weak decision trees in sequence, where each new tree tries to correct the errors made by the previous ones, together with numerical interpolation and linear regression for duplex assays, achieving 100% semiquantitative accuracy and 3.7% mean absolute quantification error [12].

A further complex curve analysis may require differentiation of true amplification from background noise. This is more complex in multiplex reactions because more than one template is amplified simultaneously. Deep domain adaptation approaches using transformer-based conditional domain adversarial networks (T-CDAN) eliminate data distribution differences between source domains (synthetic DNA) and target domains (clinical isolates), achieving 93.1% curve-level accuracy and 97.0% sample-level accuracy for carbapenem-resistant gene classification [39]. In deep domain adaptation, the systematic difference between training datasets and real-world data are overcome. T-CDAN is a domain adaptation technique that uses a neural network-based architecture and learns to ignore instrument or batch-specific variations while focusing on biologically relevant signals. Deep learning approaches thus allow the elimination of systematic variation between instruments and reagent lots, which are the main reasons for cross-laboratory variability.

Other advanced automated analysis systems have followed new approaches in curve interpretation. In this regard, the f_0% method, based on a modified flexible sigmoid fitting amplification curve with background noise subtraction, outperformed classic Cq methods in both absolute and relative quantification cases [39]. The method is based on a six-parameter model. A four-parameter sigmoid component fits the amplification curve and estimates maximum fluorescence, decay rate of efficiency, initial efficiency, and inflection cycle. A two-parameter linear component fits background noise by estimating the baseline slope and intercept. Compared to the standard Ct method, f_0% method decreased the coefficient of variation by 1.66-fold in absolute quantification and 1.76-fold in relative quantification, while variance was reduced by 2.78-fold and 3.13-fold, respectively [40].

Altogether, the two main AI frameworks across studies were: (a) tree-based models (XGBoost) trained on hand-engineered features (mean, slope, and variance ratio) for relatively simple multiplexing, and (b) deep domain adaptation (T-CDAN) for generalizing across platforms. The f_0% method offers a third, non-ML but AI-inspired approach (sigmoid fitting with background subtraction) that reduces coefficient of variation by 1.66-fold. A persistent pain point remains: no single method works across all instrument/reagent combinations; domain adaptation or retraining on local data remains necessary.

3.2. Cq value extraction and quantification

The determination of the Cq value is the primary basis for any further quantification based on qPCR data. However, Cq is not an absolute measurement due to its dependence on many experimental factors, such as the instrument, reaction efficiency, fluorescence monitoring chemistry, and several other experimental variables [41]. In this respect, ML has been instrumental in developing new methodologies to retrieve more meaningful quantitative information from the raw data provided by the Cq parameter.

The third derivative zero (TD0) limiting component quantification method, as implemented in RDML-Tools, is more independent of machine characteristics than classical Cq calculation methods and allows one to calculate Ncopy, i.e., the number of initial copies. Such a parameter is very straightforward to interpret and compare between instruments [42].

Other solutions based on deep learning approaches have been developed, such as shinyCurves, which allows users to quickly perform visual inspection of multisource qPCR data from hundreds of plates. The samples are automatically called as positive, negative, or undetermined according to user-defined criteria, and the results are graphically displayed with annotations for quick quality checking [43]. This software has been developed with a non-programmer audience in mind and provides a user-friendly and intuitive interface for analyzing qRT-PCR data.

The TD0 method (implemented in RDML-Tools) and the shinyCurves visual platform are approaches that replace fixed Cq thresholds with instrument-independent metrics like Ncopy or initial copy number. However, all these approaches, including TD0 and shinyCurves, rest on the premise of a sigmoidal shape of the amplification curve, a condition that is not met in inhibited or degraded samples. New methods of quality control using AI (e.g., outlier detection) are just emerging to address this limitation [44].

3.3. Positive, negative amplification classification

Automated classification of amplification results into positive, negative, or ambiguous categories constitutes an important application for ML. Binary threshold-based methods have conventionally been employed to call such data, but due to the continuous nature of the Cq value distributions, these approaches generally give rise to false positives and false negatives [45].

In this respect, researchers have developed probabilistic classification methods via AI frameworks based on the nature of the amplification curves, the end-point fluorescence values, the slopes, and the inflection points to achieve a superior classification performance. A typical example is the FastFinder platform for HPV genotyping. It provides both qualitative and quantitative improvements for qPCR assays through standardized software and AI-based curve interpretation [46]. This removes most of the variability caused by manual analysis while delivering clinically relevant results.

For multiplex PCR, where distinguishing several targets from fluorescence information is challenging, deep neural networks and recurrent architectures have shown great promise. For instance, one study merged deep learning with a microfluidic paper-based device to detect synthetic SARS-CoV − 2 RNA. Rather than waiting 40 cycles, three types of recurrent models, namely RNN (simple recurrent network), LSTM (long short-term memory), and GRU (gated recurrent unit) prediction, made a call after only 13 cycles. All three networks process data sequentially, much like reading a PCR curve cycle after cycle. These models, thus, recognize temporal patterns; in this case, the models learned to recognize the shape and progression of early amplification curves and predict whether the sample would eventually become positive (Fig. 4). Using this approach, test time was shortened by approximately 40 min (67.5% reduction). The system could also adaptively decide on a cycle threshold for each reaction, as opposed to using a fixed cutoff value [47].

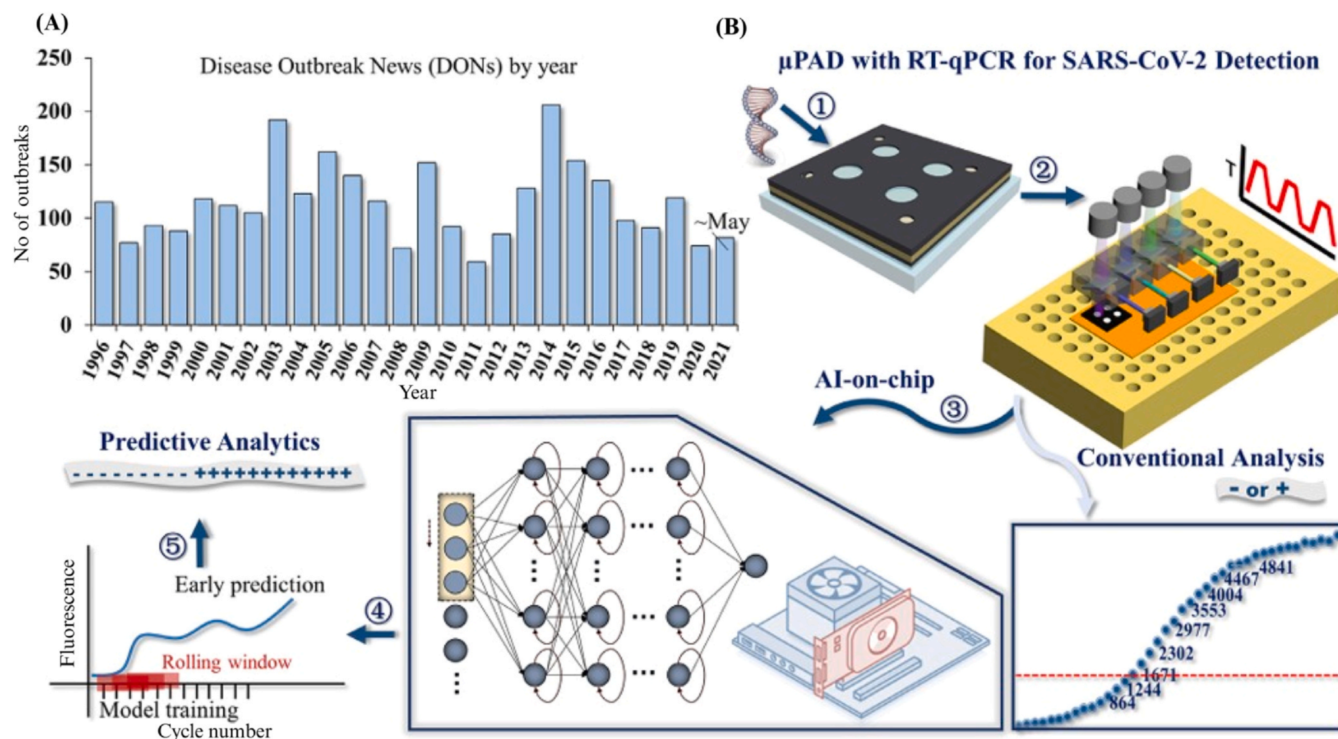


Fig. 4. Methodological framework of the AI-aided on-chip nucleic acid assay. (A) Disease outbreaks recorded by the WHO. (B) Schematic of the proposed approach. Steps 1–5 illustrate the workflow: sample collection, RT-qPCR in µPADs, model training and validation, early prediction on time series, and final output. Adopted from ref. [47].

More recently, support vector machines (SVMs), a classification technique that draws a boundary separating two groups of data, were coupled with feature extraction (condensing raw signals into summary measures that capture key aspects of the data) to classify breath samples from a COVID − 19 screening program as positive or negative [48]. The classification approach yielded an overall accuracy of 88% − 95% in detecting COVID − 19 patients, with sensitivity and specificity in the range of 86% − 94% and 88% − 95%, respectively.

Recurrent neural networks (RNNs, LSTM, and GRU) have a significant advantage over other methods in that they enable early prediction of final positivity after only 13 cycles-a 67.5% time reduction because these networks learn temporal patterns in the early amplification. By contrast, feature extraction combined with an SVM yields 88% − 95% accuracy but requires the full-length curves. The main limitation lies in interpretability. Would a clinician accept an early “positive” call made by an LSTM black box without understanding the decision’s rationale? Hence, explainable AI approaches, such as attention maps over cycles, must be investigated.

3.4. Amplification efficiency analysis and normalization

In PCR, amplification efficiency is a major source of variability. The traditional approach assumes all reactions run at exactly 100% efficiency, meaning the template doubles every cycle. In reality, efficiencies vary due to primers, template quality, inhibitors, and instrument differences. When calculations ignore this, relative expression ratios can be off by as much as 100-fold [45].

Advanced ML approaches provide predictive and corrective capabilities regarding efficiency variability resulting from factors such as primer concentration, template composition, PCR inhibitors, and instrument-specific properties. One innovative study used a BSA hydrogel to control the release of Taq polymerase, achieving more than 20 cycles of linear amplification, enabling correlation between amplification slopes and Ct values ($r = -0.9455$) with quantification of efficiency through exponential equations fitted with $R^2 = 0.9995$ [49]. This physics-informed ML approach enables absolute quantification without standard curves by reversing the relationship between initial DNA template concentration and end-point fluorescence intensity.

Touchdown qPCR (where annealing temperature gradually decreases) improves specificity, but optimizing the temperature steps is tedious. ML can now automate this optimization. In one study, AI-driven parameter optimization decreased average Cq values for reference genes by 3.7–4.7 cycles. More importantly, it enabled robust amplification at annealing temperatures below 55°C, conditions where conventional qPCR produces non-specific products. They reported a modified protocol of touchdown qPCR with a final annealing of 50°C to detect gene expression. This expands qPCR's reach to low-abundance and otherwise challenging targets [50]. Taken together, these advances significantly extend the scope of qPCR in terms of targets and experimental conditions.

In these contexts, ML has enabled two advances that would otherwise have been impossible using classical efficiency-based approaches: First, ML has enabled absolute quantification in the absence of standard curves by inverting the relationship between the initial amount of template and the endpoint fluorescence (in physics-informed ML); and second, it has enabled the automation of touchdown qPCR annealing temperature optimization, leading to decreases of 3.7–4.7 in Cq. Most efficiency models to date have been developed under "ideal" conditions (i.e., purified DNA without inhibitors). How these models will behave in clinical samples containing known inhibitors remains an open remains an open challenge, highlighting the need for future AI models trained with real-world matrix effects.

4. AI-driven PCR automation

Physical PCR workflows remain highly manual and error-prone, despite more advanced primer design and analysis of qPCR data. There are three main pain points in automation. First, sample preparation and plate handling consume up to 70% of hands-on time and are a major source of pipetting errors. Second, thermocyclers, liquid handlers, and

Laboratory Information Management Systems (LIMS) systems each store data in their proprietary formats, hindering cross-system data integration and the standardization of communication protocols. Third, manual inspection of amplification curves is required for report generation and clinical interpretation, which substantially prolongs diagnostic turnaround time in most clinical scenarios. Most of the current AI-powered automation systems tackle these issues through robotic transport of labware, standardization based on SiLA, and automated quality control with human-in-the-loop flagging.

4.1. Automated data collection and integration systems

Automation in laboratories has moved from simple liquid handling systems toward integrated platforms of robotic manipulators, optical detection, data management, and AI for decision-making. In this respect, one of the advanced systems recently described for automated qPCR measurement is qPCRBot (Fig. 5). It integrates standardized communication protocols, robotic labware transport, and automated data analysis. The system employs a SiLA 2-based client-server architecture, where SiLA (Standardization in Lab Automation) provides a universal communication standard that enables different laboratory instruments to work together seamlessly. Thanks to this architecture, the robotic manipulators and qPCR instruments coordinate automatically: the robot shuttles plates between stations while the qPCR instrument cycles autonomously. All experimental data are then exported in a standard XML format, immediately importable into LIMS and ready for analysis [51]. The standardization of experimental data acquisition, from the raw data to the level of data formatting, facilitates cross-platform data integration and advanced analytics by removing human variability.

Therefore, high-throughput PCR workflows rely on a standardized protocol for the automated extraction of nucleic acids and sample preparation. One example is the Hamilton STARlet liquid handler, which, when combined with customized barcode scanning scripts, enables extraction-free direct PCR from saline gargle samples, reducing hands-on labor time by 76% relative to manual workflows while maintaining equivalent diagnostic performance [52]. Likewise, preprocessing ThinPrep cervical specimens for automated HPV testing gave 95.6% agreement with a manual protocol for 226 samples tested, with 100% agreement on the 99 samples tested by both approaches. These findings confirm that automation did not reduce reproducibility [53].

The integration of data from PCR includes more than fluorescence measurements. It extends to quality assurance parameters such as monitoring amplification kinetics and melting characteristics, and verifying thermocycler performance. In that sense, ML is well-suited to these kinds of rich data sets. Barcode integration allows samples to be tracked through preprocessing and amplification, making the data entirely traceable. However, nucleic acid extraction automation is platform-dependent. As an example, a study comparing QIAsymphony RGQ and QIAcube revealed important differences in detection rates for cytomegalovirus (CMV) and Epstein-Barr virus (EBV). Specifically, QIAsymphony RGQ gave higher rates of detection than QIAcube for CMV (59.5% vs. 43.8%) and EBV (59.0% vs. 42.7%) [54]. These results clearly indicate that each automation platform needs to be validated separately before being used in a diagnostic workflow.

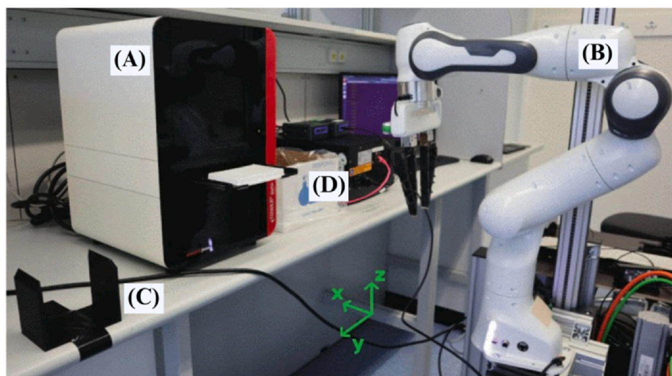


Fig. 5. Setup components of the qPCRBot (A) qPCR thermocycler with well plate on tray, (B) robotic manipulator, (C) well plate stand, and (D) trash bin. Reprinted with permission from ref. [51].

Common to all automation platforms discussed above is the insight that standardization is just as important as the robotics, whether this relates to SiLA 2, CML export or barcode tracking. In the absence of standardized data formats, even the most powerful AI models will not be scalable across laboratories. Consequently, each automation system must be validated separately before clinical use.

4.2. Microfluidics PCR automation

Parallel to these robotic liquid handlers, another leading paradigm in microfluidic PCR automation is the miniaturization and integration of multiple reaction steps onto a single chip. Recent developments fall under three broad categories. First are centrifugal microfluidic systems. These platforms use spinning discs to control fluid movement via centrifugal force. A fully automatic, high-throughput centrifugal microfluidic workstation has recently been demonstrated for rapid respiratory virus diagnostics [55]. The system integrates RNA extraction, gene amplification, and fluorescence detection on a single disc accommodating six distinct units, enabling simultaneous analysis of up to 36 targets in one run. A real-time centrifugal plastic microfluidic chip with a dual-valving strategy (wax valve + membrane valve) achieved multiplexed PCR detection within 20 min [56]. Another centrifugation-free microfluidic chip using passive capillary pressure–based valves performed nucleic acid extraction in 10 min followed by photonic qPCR amplification within 8 min, demonstrating diagnostic sensitivity comparable to conventional RT-PCR [57].

Second are integrated digital PCR (dPCR) microfluidic platforms. Commercial systems such as the QIAcuity Digital PCR System (Qiagen) and the QuantStudio Absolute Q AutoRun Suite (Thermo Fisher) now integrate partitioning, thermal cycling, and imaging on a single fully automated instrument [58]. These systems greatly reduce the hands-on time and retain the high accuracy of absolute quantification. Two new platforms under research are a digital microfluidic chip combined with a dual-mode thermal controller (4 °C to 95 °C) for the fully automatic integrated nucleic acid detection [59] and a low-cost smartphone-based digital PCR aiming at the minimization of the footprint and cost of the instrument [60].

Lastly, portable point-of-care (POC) microfluidic devices. For decentralized testing, a multifunctional portable nucleic acid detection platform has integrated both RT-qPCR and RT-LAMP within a unified microfluidic device, achieving a detection limit of 2.0 copies μL^{-1} for RT-qPCR while reducing overall size by approximately 90% compared with conventional qPCR instruments [61]. A low-cost portable PCR device (WACAN chip) integrating MEMS-fabricated micro-heaters and a liquid cold plate achieved temperature stability within ± 0.5 °C and successfully amplified bacterial DNA [62].

4.3. Integrated data analysis pipelines

Automated analysis pipelines for PCR must be efficient and rigorously scientific and fit-for-purpose in their regulatory context. Total laboratory automation (TLA) systems embrace the integration of pre-, during, and post-analytical phases, facilitating advanced real-time quality control and auto-verification, thanks to middleware integration [63].

Such an approach can be realized using an intuitive graphical interface, like Click-qPCR, which makes complex qPCR data analysis accessible to many. This application takes CSV files containing ΔCq and $\Delta\Delta Cq$ using reference genes selected by the user, performs Welch's *t*-

tests for pairwise comparisons, and conducts one-way ANOVA with Dunnett's post-hoc test for experiments involving multiple groups. Outputs include interactive visualization showing means with standard deviations and individual data points, along with publication-ready images and statistical summaries [64]. By targeting experimental biologists without, or with limited, programming skills, this technology shows one way in which well-thought-out integrated pipelines can democratize access to powerful analytical tools.

This consideration of usability is critical, given that no matter how powerful an AI model is, it is of no use if experimental biologists cannot interact with it. Thus, any future systems that implement curve-fitting AI should directly integrate it into such interfaces, without the need for separate Python environments.

4.4. Automated report generation and clinical interpretation

In this final stage, raw data produced by PCR are summarized into clinically meaningful reports suitable for patient management and epidemiological surveillance. This approach can contribute to consistent clinical application of decision rules by embedding AI interpretations within the process and ensuring that borderline or ambiguous results are flagged for manual review.

PCR.Ai (www.pcr.ai) is an automated system that can handle the manual data analysis and quality control steps routinely encountered during clinical pathogen testing by multiplex quantitative real-time PCR. The system has been tested in clinical studies to measure its accuracy and impact as the last step of final interpretation and verification of routine in-house multiplex qPCR testing for CMV, EBV, and adenovirus on blood samples. Results showed 100% concordance between PCR. AI and the current manual interpretation methods for both detection and quantification of all three viruses. Beyond accuracy, the platform delivered significant time savings, reducing hands-on time by 63 min per run [65].

This proof of principle shows that full automation of result interpretation for routine viral load tests, including CMV, EBV, and adenovirus, can be achieved without loss of accuracy relative to manual interpretation and with a time gain of 63 min per run. However, ambiguous or borderline results are still flagged for manual review. Thus, an approach that hybridizes human and AI optimizes efficiency while maintaining safety. The remaining bottleneck concerns regulatory issues; until now, no AI-based system for interpreting PCR results has been cleared by the FDA as a Class II device, and thus, in certain jurisdictions, clinical adoption is impeded.

5. Discussions

After reviewing AI applications in pre-PCR, post-PCR, and automation phases, it is essential to take stock of the broader unresolved bottlenecks within this field. These include: (1) generalization – models trained on one instrument or organism family often fail on another; (2) data scarcity – high-quality negative data (failed PCRs) are rarely published, creating survivorship bias; (3) standardization – no universally accepted benchmark dataset exists for comparing AI-PCR methods; and (4) clinical implementation – regulatory, interpretability, and reimbursement barriers slow real-world adoption. The following subsections elaborate on what can be done with the current state, what cannot, and what the way forward would look like regarding these bottlenecks.

To enable direct comparison of methodological choices and reported outcomes across studies, Table 2 [12,13,18,27,39,40,47–49,65] summarizes key technical parameters of the AI-PCR systems reviewed.

5.1. Capabilities and current applications

5.1.1. Diagnostic applications and pathogen detection

Collectively, the articles reviewed here demonstrate that the current use of AI-augmented PCR optimization dramatically improves the diagnostic utility of this assay (Table 3) [13,27,38,39,41,45,47-49,55,59]. From this perspective, the enormous diagnostic potential for infectious disease emergencies is well illustrated by the recent COVID − 19 pandemic and 2022 Mpox outbreaks. Indeed, several groups have described systems that employ deep learning to detect SARS-CoV − 2 from RT-PCR data while AI models designed for variant discrimination rely exclusively on amplification curve profiles, with sensitivities and specificities both above 97% for clinical samples [66]. In line with this, we have previously described the BioInnovate AI platform, which was adapted specifically for emerging infectious diseases, achieving a 90% reduction in assay development time while retaining high predictive accuracy for primer-probe set identification [13]. From this perspective, the diagnostic time gap between the identification of a novel pathogen and an outbreak might be reduced or, in the best-case scenario, eliminated, which would be an extremely important step forward in preparing for future outbreaks.

Apart from single-target detection, AI-driven analytical methods have also been used for multiplex detection. One study reported that the velocity-controlled PCR (VC-PCR) system, which integrates AI frameworks, could detect three different pathogens simultaneously with a qualitative accuracy of 100% [12]. From the viewpoint of clinical laboratory medicine, it is highly desirable for the clinical laboratory to detect more than one pathogenic species from one clinical specimen in one reaction to improve laboratory efficiency and reduce turnaround time for clinical decisions. Other similar systems have also been used for fungal pathogen surveillance. For example, the AltoStar system has been demonstrated to have a flexible workflow that can accommodate both commercial and laboratory-developed tests for *Candida auris* identification and provides 100% analytical specificity [67].

5.1.2. Food authentication and quality control

In the field of food authentication and quality control, where regulatory requirements are growing and consumer safety issues are on the rise, this AI-augmented PCR technology would likely find a promising application. For example, one study demonstrated that species-specific primers targeting mitochondrial D-loop regions, combined with real-time PCR, achieved reliable detection of canine DNA in goat satay, reporting a limit of detection of 31.25 pg/μL and an amplification efficiency of 109.7% [68]. Thus, from this perspective, AI-optimized assays may be useful in real-life scenarios requiring food authentication under certain cross-species contamination situations.

In addition, multiplexing capabilities obtained through AI-driven analysis would open new vistas for application in the detection of various pathogens for food safety monitoring. At this point, the combination of automated sample preparation with automated analysis and ML application will be very promising for food safety and quality control.

5.2. Clinical implementation barriers and regulatory pathways

Although these AI-enabled PCR systems seem promising from an analytical standpoint, few studies discuss the real-world challenges of implementing such systems in routine clinical diagnostics.

5.2.1. Regulatory approval

Most AI-PCR systems described in the literature are research-use only. Regulatory bodies such as the FDA (US), CE-IVD (Europe), and NMPA (China) require locked algorithms for *in vitro* diagnostic (IVD) clearance, meaning the AI model cannot continuously learn or update after deployment. This is counterintuitive to the key strength of ML, that model accuracy improves with access to more data. So far, only a handful of AI-assisted PCR platforms have gained regulatory clearance, and none are based on fully adaptive deep learning. For example, the PCR.Ai system [65] was validated as a locked algorithm after initial training on local data, and then submitted for CE-IVD marking. Thus, the developer must decide whether the product is a locked algorithm that may be more accurate but will never be regulated.

Table 2
Technical comparison of AI methodologies in PCR studies.

PCR task	AI model(s) used	Dataset description	Key performance metric	Main advantage reported	Limitation / trade-off noted	Study (first author, year)
Primer-probe efficacy prediction	LGBM	Large proprietary primer dataset (not publicly available)	AUC = 0.97	Reduces assay development time by 90%	Model not tested on highly multiplexed assays; data not open-source	Lin et al. [13], 2025
Plant virus primer design	Random Forest	Plant virus genomic sequences (NCBI)	Not explicitly reported; validation rate > 90%	Automates sequence retrieval + parameter optimisation	Limited to plant viruses; needs retraining for other taxa	Ghorbani et al. [18], 2025
Duplex SYBR Green qPCR classification	XGBoost + feature engineering	Oral microbiota samples (specific size not stated)	100% semiquantitative accuracy; 3.7% mean absolute error	Integrates velocity-controlled PCR hardware with AI	Requires specialised VC-PCR hardware; not transferable to standard thermocyclers	Long et al. [12], 2026
Single-channel multiplex curve classification	T-CDAN (Transformer + domain adversarial)	Synthetic DNA (source) + clinical isolates (target)	93.1% curve-level; 97.0% sample-level	Cross-platform generalisation without retraining	Needs paired source/target domain labels; high computational cost	Mao et al. [39], 2023
Early prediction of PCR positivity	RNN, LSTM, GRU	SARS-CoV − 2 RNA on microfluidic paper-based device	Prediction after 13 cycles (67.5% time reduction)	Adaptive cycle-threshold per reaction; no fixed cutoff	Tested only on one pathogen; requires time-series data	Sun et al. [47], 2022
COVID − 19 breath sample classification	SVM + feature extraction	Breath samples from screening program	88–95% accuracy; 86–94% sensitivity; 88–95% specificity	Non-invasive, fast (electronic nose)	Indirect PCR application (exhaled breath, not direct amplification)	Nurputra et al. [48], 2022
Cq determination / absolute quantification	f0% method (sigmoid fitting + background subtraction)	Not specified (simulated + experimental qPCR data)	CV reduced 1.66 × (absolute) and 1.76 × (relative)	More instrument-independent than classical Cq	Non-ML (curve-fitting); still assumes sigmoidal shape	Gamal & Ibrahim [40], 2024
Automated result interpretation (multiplex qPCR)	PCR.Ai (proprietary AI)	Clinical blood samples for CMV, EBV, adenovirus	100% concordance with manual; 63 min saved/run	Fully integrated into lab workflow; reduces hands-on time	Requires initial validation on local data; not open-source	MacLean & Gunson [65], 2024
Absolute quantification without standard curves	Physics-informed ML (velocity-controlled PCR)	Purified DNA templates	Correlation r = –0.9455; R^2 = 0.9995	Eliminates need for standard curves	Requires velocity-controlled PCR hardware; tested only on purified DNA	Jiang et al. [49], 2023
Protein Tm prediction (thermostability)	ANN + MLP (ensemble)	Non-redundant protein sequences + language model embeddings	Correlation 0.89 between predicted and measured Tm	Applicable to enzyme design for PCR (e.g., heat-stable polymerase)	Not directly PCR primer Tm; indirect relevance	Tijare et al. [27], 2025

Table 3
Comparison of traditional and AI-aided PCR technologies across the workflow.

Workflow Stage	Traditional Approach	AI-Aided Approach	Key Advantages of AI-Aided	Current Limitations
Primer Design	Heuristic rules + nearest-neighbour thermodynamics; manual BLAST for off-target checking (Primer3, Primer-BLAST)	Machine learning (LGBM, Random Forest) trained on validated primer datasets; automated off-target prediction	Reduces assay development time by up to 90% [13]; handles non-linear interactions (GC% + hairpin)	Needs large, high-quality training datasets (failed primer data rarely published)
Tm Prediction	Nearest-neighbour model (ΔH, ΔS); GC% + length methods	Deep learning on modified nucleotides (AmNA gapmers); ANN/MLP ensembles for protein thermostability	Predicts Tm for artificial/modified nucleic acids; correlation up to 0.89 [27]	Requires experimentally validated Tm data for training; extrapolation risky for novel modifications
Secondary Structure Prediction	Thermodynamic free energy minimisation (mfold, OligoAnalyzer)	Deep learning (CNNs, Transformers) for RNA secondary structure	Higher accuracy on known sequence families	Generalisation crisis: fails on sequences distant from training set [59]
Amplification Curve Analysis	Fixed threshold; first-derivative maximum (FDM); sigmoidal fitting	XGBoost on engineered features; T-CDAN domain adaptation; RNN/LSTM early prediction	Reduces coefficient of variation 1.66 × [39]; generalises across instruments (T-CDAN); predicts after 13 cycles (67.5% time reduction) [38,45]	Requires high-quality curve data; domain labels for adaptation
Cq Determination	Manual or semi-manual threshold setting; 2nd derivative max	TD0 method; f0% method (sigmoid + background subtraction); automated Cq extraction (RDML-Tools)	More instrument-independent; computes initial copy number (Ncopy) [41]	Assumes sigmoidal shape; fails for inhibited/degraded samples
Positive/Negative Classification	Binary Cq cutoff (e.g., $Ct < 35$)	SVM + feature extraction; LSTM probabilistic classification	Reduces false positives near LOD; adaptive per reaction [45]	Interpretability: clinicians hesitant to trust black-box calls
Amplification Efficiency Normalisation	Assumes 100% efficiency; standard curves	Physics-informed ML; velocity-controlled PCR	Absolute quantification without standard curves; Cq reduction 3.7–4.7 cycles [47,48]	Models trained on purified DNA may fail on clinical samples with inhibitors
Data Analysis Automation	Manual plate review; Excel-based Cq export	Fully automated platforms (PCR.Ai, qPCRBot) with SiLA − 2 standardisation	63 min hands-on time saving per run [55]; 100% concordance with manual	Requires local validation and retraining; regulatory hurdles
Instrument Integration	Proprietary formats; manual data transfer	SiLA − 2 standardised communication; barcode tracking; LIMS integration [49]	Cross-platform compatibility; traceable data	Platform-dependent validation required; high initial setup cost

5.2.2. Workflow integration

Even if an AI model were flawless, any gains in accuracy would be rendered useless if the AI model could not be integrated into a laboratory's existing LIMS or daily routine. The qPCRBot [51] addresses this by using SiLA − 2-standardized communication, but most AI tools are offered as standalone Python scripts or cloud services. Clinical laboratories are reluctant to install third-party software due to the validation burden, cybersecurity policies, and a lack of IT personnel. For AI to be successfully implemented in routine clinical diagnostics, the output from the AI system, whether the called Cq values or positive/negative flags, needs to be automatically read by the laboratory's middleware without the need for manual entry.

5.2.3. Reimbursement and health economics

No cost-effectiveness study has yet been published for any AI-enabled PCR system. Time savings, such as 63 min per run [65], and reduced hands-on steps, such as 76% [52], are appealing metrics, but payers, including private insurers and national health systems, will want evidence that an AI improves patient outcomes or reduces the overall cost of care rather than merely improving efficiency within a laboratory. Future prospective studies should address diagnostic yield, time to clinical decision, and the avoidance of unnecessary follow-up tests.

5.2.4. Clinical trust and interpretability

Laboratory directors and clinicians want to know why a result is positive or negative. A black-box AI model, particularly deep neural networks, offers no such explanation. Therefore, accepting AI calls on borderline samples may be challenging. Some groups have addressed this by applying explainable AI methods such as SHAP or feature importance plots [38]. Future AI-PCR systems should be designed to be as interpretable as possible; for example, they should show the user which cycle or feature drove the classification decision.

5.2.5. Data privacy and decentralized learning

Because patient PCR data contains genetic information, it cannot be uploaded to a cloud-based AI service under the current GDPR (Europe) and HIPAA (USA) guidelines [69]. Federated learning, where the AI model is sent to the local server of each laboratory, trained locally, and only anonymized updates to the model are sent back, is an emerging solution. Although no PCR application has been reported yet, federated learning has been applied to several diagnostic imaging tasks and enables collaborative AI training from multiple hospitals without sharing raw patient data [70,71].

In summary, although AI-enabled PCR has performed very well in retrospective analyses of archived samples, prospective validation remains limited in real-time clinical workflows. This will require further emphasis on addressing regulatory, integration, economic, interpretability, and privacy issues in addition to further improvements in algorithmic accuracy.

5.3. Limitations and future perspectives

5.3.1. Generalization challenges and data-driven limitations

Despite remarkable advances, significant limitations persist in applying AI to PCR analysis. The most profound one is the so-called generalization crisis, in which deep learning models have been applied to the prediction of nucleic acid structures. While such models generalize well within the training domain for sequence families used during training, they collapse when presented with new sequences. Indeed, deep learning models for RNA secondary structure prediction suffer from rapid performance degradation with increasing sequence dissimilarity from the training set. Moreover, across all learning-based models, in-distribution performance is inversely correlated with out-of-distribution performance [32]. Thus, these considerations must be taken into account in systems for designing PCR primers trained on

particular organism groups or specific sequence contexts.

Specificity in primer design is still a significant problem despite ML approaches. Primers designed using such *in silico* prediction methodologies were found to be specific; however, when tested in the laboratory, they produced non-specific amplifications. This disparity between predictions and experiments indicates that many factors affecting PCR specificity, especially in complex sample matrices, are still poorly understood. For instance, a recent study evaluated primers for environmental DNA (eDNA) surveillance and found that approximately 90% of primer pairs that appeared specific during BLAST-based in silico analyses yielded non-specific amplifications during experimental testing [72]. Hence, despite advances *in silico* approaches, experimental validation will continue to be necessary.

A recent problem was highlighted related to the quality of the genome sequence used for designing primers [73]. In this particular case, primers designed for the detection of monkeypox overlapped with poorly sequenced and poorly assembled regions of MPXV genomes for several lineages, stressing the need for iterative primer design with quality control at the sequence database level and with refinements in genome sequence assemblies.

Data scarcity is another significant limitation, along with the lack of well-defined, standardized, and high-confidence negative data. For example, in this domain, an AI model may be trained on relatively small datasets, such as the 316 primer pairs dataset that is publicly available. Although this dataset is valuable, it is limited when considering the high-dimensional nature of the PCR biochemistry space [74]. Moreover, the literature is subjected to a high level of publication bias, as failed PCR experiments are hardly ever published, resulting in a “survivorship bias” in which models learn an unrealistic distribution of primer success [75]. Parameters such as the number of predicted primer-binding sites in complex genomes, DNA secondary structure, and buffer composition are seldom uniformly encoded in training datasets but are often critical to explaining PCR failure [75].

5.3.2. Standardization and cross-platform reproducibility

Challenges in standardization have restricted the applicability of AI-driven PCR systems in different laboratories on different instruments. Many authors have worked on making their methods less dependent on the machine. Still, at a fundamental level, differences in the optical system, thermal cycling, and reagent formulation mean that adding more training data does not solve the problem. For example, Cq values obtained from identical samples and primers on different instruments routinely differ by 1–3 cycles, rendering machine-dependent threshold approaches problematic [45]. This variability violates the assumption of data portability since AI models require similar input features.

Aside from well-to-well variation, the use of AI for digital PCR faces significant challenges concerning the generalizability of the models across different microfluidic and imaging platforms, where a change in the characteristics of droplets or chambers necessitates significant retraining or domain adaptation to avoid performance degradation [76,77].

5.4. Future directions and emerging technologies

Another avenue that holds promise to further improve AI-augmented PCR systems is the emerging field of physics-informed ML. Indeed, recent studies have shown that embedding domain knowledge of thermodynamic stability, secondary structure formation, and PCR kinetics within deep neural network architectures increases the performance and generalizability of ML models. For example, there have been deep learning approaches aware of structural ability in which free energy calculations have been embedded in the loss functions, showing higher performance than entirely data-driven approaches [78]. Following this line of thinking, models for primer design and amplification curve analysis could be further developed by embedding biochemical constraints.

Several recent developments in sample preparation and analysis could also benefit from further integration with AI technologies. For example, while LAMP has several application-specific advantages over conventional PCR, the recently developed automated SMART-LAMP system has demonstrated a throughput of 40,000 samples per day in a fully scalable workflow. The recently developed automated SMART-LAMP system has demonstrated a throughput of 40,000 samples per day in a fully scalable workflow [79]. Future efforts in AI-powered data analysis on LAMP platforms or hybrid amplification systems, which take advantage of the benefits of several amplification systems, are among the critical future directions. Despite the promise that novel AI techniques hold, so far, in clinical and laboratory settings, the best pragmatic approach remains interpretability-oriented, user-interactive human-AI systems. Fully autonomous solutions may not be the answer; rather, interactive frameworks allowing domain expertise and contextual information from scientists and clinicians to guide AI analyses hold the most promise for translation into real-world settings [80]. Beyond interactive AI frameworks, another near-term technological goal is the commercialization of highly robust qPCRBots that can automate experimental repetition for data integrity verification. Such systems would liberate laboratory staff from tedious manual replication tasks while improving the consistency of training data for downstream AI models. Nevertheless, a critical bottleneck is that training these robots to handle diverse experiment scenarios (e.g., varying sample matrices, reagent lots) require extensive and expensive empirical work. Consequently, future research should prioritize cost-effective training paradigms, such as transfer learning from simulated environments or collaborative benchmarking among multiple laboratories, to accelerate the deployment of reliable qPCR automation. These would be the approaches where the capabilities of AI in rapidly analyzing data and identifying complex patterns are harnessed while avoiding the pitfalls of complete automation.

To date, AI-PCR systems have used only one data type, the fluorescence amplification curve. In reality, clinicians synthesize multiple data types, such as clinical symptoms, epidemiological data, prior diagnostic results, and sometimes imaging and genomic sequencing, to arrive at a final diagnosis. Combining data from different sources and modalities in a single model (Multimodal AI) has been shown in principle to improve diagnostic accuracy and clinical utility. For example, a multimodal model could integrate (i) RT-PCR Cq values, (ii) patient symptom onset day, (iii) viral lineage from sequencing, and (iv) chest CT imaging features to predict disease severity or treatment response. Multimodal fusion models outperform their unimodal counterparts for a variety of fusion scenarios as demonstrated by preliminary proof-of-concept studies outside the PCR domain [81,82]. Key challenges include aligning data collected at very different sampling rates, the frequent absence of some modalities in real-world settings, and increased model complexity that requires larger training datasets [81,83].

The development of Lightweight AI is critical for deploying ML models directly onto portable, point-of-care PCR devices, which often operate with limited computational resources, memory, and battery power [84]. This area is active in research through different model pruning techniques, quantization, and knowledge distillation. In the best case, these techniques have allowed deep learning models to be compressed 10–100 times with only minor degradation in performance [85]. Although this has not yet been demonstrated for PCR on the device, the basic proof of concept has been shown in other medical domains [84].

dPCR allows for the absolute quantification of nucleic acids by partitioning a sample into thousands of individual reactions and, traditionally, counting the number of positive partitions [86]. However, manual image analysis can be slow and error-prone. In this context, the SAM-dPCR method is currently the most advanced. It uses a “foundation model” that has been previously trained on millions of general-purpose images to quantify dPCR droplets in zero-shot mode [86]. This method does not require further training on new data, can analyze

images in less than 4 s with an accuracy greater than 97%, and can be applied to various dPCR platforms. These powerful AI tools are rapidly moving dPCR towards a smarter, automated, and clinically more relevant diagnostic technology [87,88].

Finally, it remains critical to focus on defining standard practices for AI model training and validation. International cooperation will be necessary to define benchmark datasets and standard validation procedures to allow for the rapid yet responsible implementation of AI technologies across different laboratories. Given the growing penetration of AI-based systems in diagnostics and research, the urgent need is to develop international standards similar to those established for standard PCR data analysis.

In summary, the four major interconnected pain points – generalization, data quality, standardization, and clinical translation – outline a complex picture. Models that perform well on in-distribution test sets collapse when sequence similarity drops below 70% [32]. Training datasets lack systematically annotated negative examples [75], and no two laboratories use identical PCR protocols, making cross-study comparison difficult. Moving forward, the community needs to agree on open benchmark datasets (including failed PCRs), embrace domain adaptation techniques, and establish regulatory pathways for locked versus continuously learning AI models.

6. Conclusion

The integration of AI and ML into PCR technology marks a radical shift in how molecular biology is practiced. From smart primer design and advanced data analysis to automated workflows, AI has overcome long-standing methodical obstacles and unveiled previously unavailable capabilities. Indeed, ML algorithms have thus far been very successful in modeling PCR outcomes, designing target-specific primers and probes, deciphering complex amplification curve patterns, and integrating diverse data modalities to support clinical decisions.

Yet, for all its promise, the rise of AI in PCR does not diminish the enduring value of human reasoning and hands-on laboratory work. On the contrary, the most successful implementations are likely to be those in which machines augment rather than replace human expertise. Experienced researchers and clinicians remain essential for framing biological questions, interpreting ambiguous results, identifying subtle anomalies that fall outside training distributions, and exercising judgments in contexts where automation alone would be insufficient. The physical act of experimental design, sample handling, quality control, and troubleshooting, rooted in tacit knowledge accumulated through practice, cannot be fully encoded into algorithms. Thus, this field needs to be cautious about over-automating away the very human expertise upon which good science depends.

That said, many issues remain unresolved: limited generalizability, the need for large amounts of data, difficulties in standardization, and regulatory questions that will need to be addressed through continued research and development. In the future, AI-powered PCR will most likely be hybrid approaches that leverage the interpretability and domain grounding of traditional methodologies with the pattern recognition and optimization capabilities of ML. As these new technologies mature and become further embedded in clinical and research workflows, continuous efforts will be needed to ensure validation, standardization, and regulatory compliance. Ultimately, if AI is ever to fulfill its promise of significantly improving human health and expanding our understanding of the natural world, it will need better algorithms and, at the same time, a strong, thoughtful human presence behind the laboratory bench.

Data Sources and Note

This comprehensive literature review synthesizes findings from 80 + peer-reviewed publications. No synthetic data visualizations were generated, as the review comprehensively synthesizes existing methodological advances and comparative analyses from published literature.

CRediT authorship contribution statement

Stephanie Andaluz: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Aman Lv:** Data curation. **Pengyu Yu:** Data curation. **Wenhao Hui:** Data curation. **Ren Shen:** Supervision, Writing Review, Funding acquisition. **Tzu-Ming Liu:** Supervision, Writing – review & editing. **Miguel Brito:** Supervision, Writing – review & editing. **Pui-In Mak:** Supervision, Writing – review & editing. **Rui P. Martins:** Supervision, Writing – review & editing. **Yanwei Jia:** Writing – review & editing, Validation, Supervision, Methodology, Funding acquisition, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by the Macau Science and Technology Development Fund (FDCT) [FDCT0168/2023/RIA3, FDCT0001/2025/RID, FDCT0213/2024/AGJ, FDCT0001/2025/NRP]; University of Macau [UMDF-TISF/2025/014/IME, SRG2024–00057-IME].

References

[1] H. Zhu, H. Zhang, Y. Xu, et al., PCR past, present and future, BioTechniques 69 (2020) 317–325, https://doi.org/10.2144/btn-2020-0057.
[2] C.T. Wittwer, A.C. Hemmert, J.O. Kent, et al., DNA melting analysis, Mol. Asp. Med. 97 (2024) 101268, https://doi.org/10.1016/j.mam.2024.101268.
[3] J. de Korne-Elenbaas, L. Caduff, A. Lison, et al., Design, validation, and implementation of multiplex digital PCR assays for simultaneous quantification of multiple targets, Lett. Appl. Microbiol. 78 (2025) ovae137, https://doi.org/10.1093/lambio/ovae137.
[4] M. Al-Remawi, A.S.A. Ali Agha, F. Al-Akayleh, et al., Artificial intelligence and machine learning techniques for suicide prediction: integrating dietary patterns and environmental contaminants, Heliyon 10 (2024) e40925, https://doi.org/10.1016/j.heliyon.2024.e40925.
[5] J. Ukwaththa, S. Herath, D.P.P. Meddage, A review of machine learning (ML) and explainable artificial intelligence (XAI) methods in additive manufacturing (3D Printing), Mater. Today Commun. 41 (2024) 110294, https://doi.org/10.1016/j.mtcomm.2024.110294.
[6] N. Doreswamy, L. Horstmanshof, Generative AI decision-making attributes in complex health services: a rapid review, Cureus (2025), https://doi.org/10.7759/cureus.78257.
[7] R.S. Aal E Ali, J. Meng, M.E.I. Khan, et al., Machine learning advancements in organic synthesis: a focused exploration of artificial intelligence applications in chemistry, Artif. Intell. Chem. 2 (2024) 100049, https://doi.org/10.1016/j.aichem.2024.100049.
[8] G. Baskar, S. Nashath Omer, P. Saravanan, et al., Status and future trends in wastewater management strategies using artificial intelligence and machine learning techniques, Chemosphere 362 (2024) 142477, https://doi.org/10.1016/j.chemosphere.2024.142477.
[9] S. Hosseini, H. Seilani, The role of agentic AI in shaping a smart future: a systematic review, Array 26 (2025) 100399, https://doi.org/10.1016/j.array.2025.100399.
[10] C.M. Hussain, G. Hussain, R. Keçili, Smart analytical chemistry: integrating green, sustainable, white and AI-driven approaches in modern analysis, TrAC. Trends Anal. Chem. 191 (2025) 118295, https://doi.org/10.1016/j.trac.2025.118295.
[11] Z. Chen, L. Xu, H. Zheng, et al., Evolution and prospects of foundation models: from large language models to large multimodal models, Comput. Mater. Contin. 80 (2024) 1753–1808, https://doi.org/10.32604/cmc.2024.052618.
[12] A. Long, W. Chen, S. Zhuo, et al., AI-assisted multiplex SYBR green I-based qPCR for the identification and quantitative analysis of oral microbiota, Anal. Chem. 98 (2026) 3177–3188, https://doi.org/10.1021/acs.analchem.5c06735.
[13] H.H. Lin, H.Y. Chung, T.H. Lin, et al., BioInnovate AI: a machine learning platform for rapid PCR assay design in emerging infectious disease diagnostics, Diagnostics 15 (2025) 1445, https://doi.org/10.3390/diagnostics15121445.
[14] D. Kowald, S. Scher, V. Pammer-Schindler, et al., Establishing and evaluating trustworthy AI: overview and research challenges, Front. Big Data 7 (2024) 1467222, https://doi.org/10.3389/fdata.2024.1467222.
[15] P. Radanliev, AI ethics: integrating transparency, fairness, and privacy in AI development, Appl. Artif. Intell. 39 (2025) 2463722, https://doi.org/10.1080/08839514.2025.2463722.
[16] Z.M. Hein, D. Guruparan, B. Okunsai, et al., AI and machine learning in biology: from genes to proteins, Biology 14 (2025) 1453, https://doi.org/10.3390/biology14101453.

[17] I.Z. Waziri, M.I. Usman, Z.Ali Dandalma, PyPCRtool: a Python package for in silico PCR and primer verification, Dutse J. Pure Appl. Sci. 10 (2024) 271–277, https://doi.org/10.4314/dujopas.v10i3a.25.

[18] A. Ghorbani, M. Rostami, E. Ashrafi-Dehkordi, et al., AutoPVPrimer: a comprehensive AI-Enhanced pipeline for efficient plant virus primer design and assessment, PLoS. One 20 (2025) e0317918, https://doi.org/10.1371/journal.pone.0317918.

[19] T. Koressaar, M. Remm, Enhancements and modifications of primer design program Primer3, Bioinformatics 23 (2007) 1289–1291, https://doi.org/10.1093/bioinformatics/btm091.

[20] M.J. Lozano, D. Dallachiesa, G. Beihammer, et al., ColabPCR: a validated Google colaboratory notebook for reproducible and precise primer design, Comput. Biol. Chem. 123 (2026) 109035, https://doi.org/10.1016/j.compbiolchem.2026.109035.

[21] S. Arvidsson, M. Kwasniewski, D.M. Riaño-Pachón, et al., QuantPrime–a flexible tool for reliable high-throughput primer design for quantitative PCR, BMC Bioinform. 9 (2008) 465, https://doi.org/10.1186/1471-2105-9-465.

[22] P.K. Singh, M.S. Donnenberg, Revolutionizing Molecular cloning: introducing FastCloneAssist, a Streamlined Python tool for optimizing primer design in restriction & ligation-independent PCR cloning, PLoS. One 20 (2025) e0306950, https://doi.org/10.1371/journal.pone.0306950.

[23] J. Guo, J. Grojean, H. Guo, ARMSprimer3: an open-source primer design Python program for amplification refractory mutation system PCR (ARMS-PCR), J. Pathol. Inform. 17 (2025) 100442, https://doi.org/10.1016/j.jpi.2025.100442.

[24] J.W. Pitsch, S.A. Wirth, N.T. Costantino, et al., CREPE (CREate primers and evaluate): a computational tool for large-scale primer design and specificity analysis, Genes 16 (2025) 1062, https://doi.org/10.3390/genes16091062.

[25] Y. Zhou, S. Ha, Y. Xu, et al., Establishment of a simple prediction method for DNA melting temperature: high-resolution melting curve analysis of PCR products, PLoS. One 20 (2025) e0321885, https://doi.org/10.1371/journal.pone.0321885.

[26] M. Kuroda, Y. Kasahara, M. Hirose, et al., Construction of a Tm-value prediction model and molecular dynamics study of AmNA-containing gapmer antisense oligonucleotide, Mol. Ther. Nucleic Acids 35 (2024) 102272, https://doi.org/10.1016/j.omtn.2024.102272.

[27] P. Tijare, N. Kumar, G.P.S. Raghava, Prediction and design of thermostable proteins with a desired melting temperature, Sci. Rep. 15 (2025) 16683, https://doi.org/10.1038/s41598-025-98667-9.

[28] A. Khaira, A. Achyar, Z. Zulyusri, et al., Primer design and optimization of annealing temperature for analysis of glutathione reductase gene expression in rice (Oryza sativa L.), J. Biol. Sci. Technol. Man. 5 (2023) 142–148, https://doi.org/10.5614/3bio.2023.5.1.3.

[29] S.C. Baek, B. Kim, H. Jang, et al., Structural atlas of human primary microRNAs generated by SHAPE-MaP, Mol. Cell. 84 (2024) 1158–1172.e6, https://doi.org/10.1016/j.molcel.2024.02.005.

[30] A.C. Chack, E.B. Harrison, C.J. Aufgebauer, et al., Abstract A023: RNA secondary structures mediate circular RNA stability and function, Mol. Cancer Ther. 23 (2024) A023, https://doi.org/10.1158/1538-8514.rnadrivers24-a023.

[31] I.V. Kutyavin, New approach to real-time nucleic acids detection: folding polymerase chain reaction amplicons into a secondary structure to improve cleavage of Förster resonance energy transfer probes in 5'-nuclease assays, Nucleic Acids Res. 38 (2010) e29, https://doi.org/10.1093/nar/gkp1138.

[32] X. Qiu, Sequence similarity governs generalizability of de novo deep learning models for RNA secondary structure prediction, PLoS. Comput. Biol. 19 (2023) e1011047, https://doi.org/10.1371/journal.pcbi.1011047.

[33] M. Saeung, J. Pengon, C. Pethrak, et al., Dirus complex species identification PCR (DiCSIP) improves the identification of Anopheles dirus complex from the Greater Mekong Subregion, Parasites Vectors 17 (2024) 260, https://doi.org/10.1186/s13071-024-06321-6.

[34] R.M. Gao, M.O. Duceppe, M. Kang, et al., Comprehensive genomic analysis of the potential limitations of several published PCR primers targeting prfA-virulence gene cluster in Listeria species, Lett. Appl. Microbiol. 76 (2023) ovad063, https://doi.org/10.1093/lambio/ovad063.

[35] C.L. Nobles, S. Reddy, J. Salas-McKee, et al., iGUIDE: an improved pipeline for analyzing CRISPR cleavage specificity, Genome Biol. 20 (2019) 14, https://doi.org/10.1186/s13059-019-1625-3.

[36] J.R. Dobosy, S.D. Rose, K.R. Beltz, et al., RNase H-dependent PCR (rhPCR): improved specificity and single nucleotide polymorphism detection using blocked cleavable primers, BMC Biotechnol. 11 (2011) 80, https://doi.org/10.1186/1472-6750-11-80.

[37] S. Jorapur, A. Srivastava, S. Kulkarni, Evaluating the usefulness of a large language model as a wholesome tool for de novo polymerase chain reaction (PCR) primer design, Cureus (2023), https://doi.org/10.7759/cureus.47711.

[38] B.J. Brintz, D.J. Operario, D.G. Brown, et al., Automated post-run analysis of arrayed quantitative PCR amplification curves using machine learning, Gates Open. Res. 9 (2025) 1, https://doi.org/10.12688/gatesopenres.16313.1.

[39] Y. Mao, K. Xu, L. Miglietta, et al., Deep domain adaptation enhances amplification curve analysis for single-channel multiplexing in real-time PCR, IEEE J. Biomed. Health Inform. 27 (2023) 3093–3103, https://doi.org/10.1109/jbhi.2023.3257727.

[40] M. Gamal, M.A. Ibrahim, Introducing the f0% method: a reliable and accurate approach for qPCR analysis, BMC Bioinform 25 (2024) 17, https://doi.org/10.1186/s12859-024-05630-y.

[41] R. Punchoo, S. Bhoora, A. Bangalee, Laboratory Considerations for Reporting Cycle Threshold Value in COVID-19, EJIFCC 33 (2) (Aug. 2022) 80–93.

[42] A. Untergasser, Q.D. Gunst, V. Benes, et al., Implementation and validation of a limiting component quantification method for qPCR, Int. J. Mol. Sci. 27 (2026) 2717, https://doi.org/10.3390/ijms27062717.

[43] S. Olaechea-Lázaro, I. García-Santisteban, J.R. Pineda, et al., shinyCurves, a shiny web application to analyse multisource qPCR amplification data: a COVID-19 case study, BMC Bioinform 22 (2021) 476, https://doi.org/10.1186/s12859-021-04392-1.

[44] L. Miglietta, K. Xu, P. Chhaya, et al., Adaptive filtering framework to remove nonspecific and low-efficiency reactions in multiplex digital PCR based on sigmoidal trends, Anal. Chem. 94 (2022) 14159–14168, https://doi.org/10.1021/acs.analchem.2c01883.

[45] A. Ruiz-Villalba, J.M. Ruijter, M.J.B. van den Hoff, Use and misuse of cq in qPCR data analysis and reporting, Life 11 (2021) 496, https://doi.org/10.3390/life11060496.

[46] A.R. Pereira, N. Redzic, S. Van Vooren, et al., Development, validation, and implementation of an augmented multiwell, multitarget quantitative PCR for the analysis of human papillomavirus genotyping through software automation, data science, and artificial intelligence, J. Mol. Diagn. 26 (2024) 781–791, https://doi.org/10.1016/j.jmoldx.2024.05.012.

[47] H. Sun, L. Xiong, Y. Huang, et al., AI-aided on-chip nucleic acid assay for smart diagnosis of infectious disease, Fundam. Res. 2 (2022) 476–486, https://doi.org/10.1016/j.fmre.2021.12.005.

[48] D.K. Nurputra, A. Kusumaatmaja, M.S. Hakim, et al., Fast and noninvasive electronic nose for sniffing out COVID-19 based on exhaled breath-print recognition, npj Digit. Med. 5 (2022) 115, https://doi.org/10.1038/s41746-022-00661-2.

[49] X. Jiang, K. Li, B. Xie, et al., An amplification velocity-controlled PCR device for accurately detecting the initial content of target DNA templates, Chem. Eng. J. 456 (2023) 141123, https://doi.org/10.1016/j.cej.2022.141123.

[50] K.A. Rahmi, H. Khotimah, M.S. Rohman, Utilization of modified touchdown qPCR to enhance sensitivity and specificity for genes with restricted expression and low optimum temperature primers, Biomed. Biotechnol. Res. J. 8 (2024) 231–237, https://doi.org/10.4103/bbrj.bbrj_111_24.

[51] H. Zwirnmann, M. Eckhoff, D. Knobbe, et al., The qPCRBot: combining automated data handling, standardization, and robotic labware transport for better qPCR measurements, in: 2025 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2025, pp. 15858–15864, https://doi.org/10.1109/icra55743.2025.11128567.

[52] V.J. Gadkar, D.M. Goldfarb, G.N. Al-Rawahi, et al., Extraction-free clinical detection of SARS-CoV-2 virus from saline gargle samples using Hamilton STARlet liquid handler, Sci. Rep. 13 (2023) 4241, https://doi.org/10.1038/s41598-023-30993-2.

[53] D. Barbieri, S. Venturoli, S. Costa, et al., Improving laboratory efficiency by automation of preanalytic processing of ThinPrep specimens for real-time PCR high-risk HPV testing, SLAS Technol. 21 (2016) 432–438, https://doi.org/10.1177/2211068215569347.

[54] H. Kim, M. Hur, J.Y. Kim, et al., Automated nucleic acid extraction systems for detecting cytomegalovirus and Epstein-Barr virus using real-time PCR: a comparison study between the QIAsymphony RGQ and QIAcube systems, Ann. Lab. Med. 37 (2017) 129–136, https://doi.org/10.3343/alm.2017.37.2.129.

[55] V.M. Phan, M.Y. Lee, H.Van Nguyen, et al., A fully automatic, high-throughput, all-in-one centrifugal microfluidic workstation integrated with a Switching-PCR assay for rapid respiratory virus diagnostics, Chem. Eng. J. 530 (2026) 173227, https://doi.org/10.1016/j.cej.2026.173227.

[56] Y. Zhang, Y. Zeng, L. Liu, et al., A real-time centrifugal microfluidic chip with dual-valving strategy for multiplexed PCR detection at point-of-care testing, Chemosensors 14 (2026) 118, https://doi.org/10.3390/chemosensors14050118.

[57] J.H. Oh, K.H. Kim, S.E. Seo, et al., Centrifugation-free sequential-flow microfluidic chip with a passive valve for nucleic acid extraction and photonic PCR verification, Chem. Eng. J. (2026) 177841, https://doi.org/10.1016/j.cej.2026.177841.

[58] “QIAcuity Digital PCR System.” Accessed: Jun. 09, 2026. [Online]. Available: ⟨https://www.qiagen.com/us/products/instruments-and-automation/pcr-instruments/qiacuity-digital-pcr-system⟩.

[59] J. Liao, C. Li, J. Liu, et al., Automatic and integrated detection of nucleic acid by using a dual-mode thermal controlled digital microfluidic chip, Anal. Chim. Acta 1334 (2025) 343415, https://doi.org/10.1016/j.aca.2024.343415.

[60] H. Zhang, X. Liu, X. Wang, et al., SPEED: an integrated, smartphone-operated, handheld digital PCR Device for point-of-care testing, Microsyst. Nanoeng. 10 (2024) 62, https://doi.org/10.1038/s41378-024-00689-2.

[61] B. Zhang, X. Li, M. Huang, et al., A portable dual-mode microfluidic device integrating RT-qPCR and RT-LAMP for rapid nucleic acid detection in point-of-care testing, Biosensors 16 (2026) 51, https://doi.org/10.3390/bios16010051.

[62] J.J. Chen, Y.J. Wang, M.A. Tsai, et al., WACAN chip: a MEMS-based microfluidic PCR platform with integrated micro-heaters and liquid cold plate for efficient nucleic acid amplification, Microsyst. Technol. 31 (2025) 2943–2960, https://doi.org/10.1007/s00542-025-05917-x.

[63] A.S. Brown, T. Badrick, The next wave of innovation in laboratory automation: systems for auto-verification, quality control and specimen quality assurance, Clin. Chem. Lab. Med. CCLM 61 (2023) 37–43, https://doi.org/10.1515/cclm-2022-0409.

[64] A. Kubota, A. Tajima, Click-qPCR: a simple tool for interactive qPCR data analysis, BIO-Protoc. 15 (2025), https://doi.org/10.21769/bioprotoc.5513.

[65] A.R. MacLean, R. Gunson, Automation and standardisation of a quantitative multiplex PCR assay using PCR.Ai, J. Virol. Meth 329 (2024) 114981, https://doi.org/10.1016/j.jviromet.2024.114981.

[66] T.Ž. Čretnik, M. Retelj, S. Janežič, et al., Accuracy of Allplex SARS-CoV-2 assay amplification curve analysis for the detection of SARS-CoV-2 variant Alpha, Future Microbiol. 17 (2022) 1125–1131, https://doi.org/10.2217/fmb-2021-0288.

[67] C.K. Lee, J.W.S. Cheng, N.A. Liyana, et al., Evaluating a flexible workflow for Candida auris surveillance on the AltoStar automation system AM16 with commercial and in-house real-time PCR assays, Ann. Lab. Med. (2026), https://doi.org/10.3343/alm.2025.0574.

[68] L.A. Lestari, W.N. Aini, M.D. Laksitorini, et al., Application of real-time PCR for analysis canine meat (Canis lupus familiaris) in goat's satay for halal authentication study, Open. Vet. J. 15 (2025) 456–464, https://doi.org/10.5455/OVJ.2025.v15.i1.41.
[69] A. Ali, V. Snášel, J. Platoš, Health-FedNet: a privacy-preserving federated learning framework for scalable and secure healthcare analytics, Results Eng. 27 (2025) 106484, https://doi.org/10.1016/j.rineng.2025.106484.
[70] S. Kempter, J.O. Brunner, F. Hanses, et al., Enhancing pandemic surveillance and testing: a simulation modeling study utilizing German multicenter data with federated machine learning, Heal. Care Manag. Sci. 29 (2026) 12, https://doi.org/10.1007/s10729-025-09752-4.
[71] S. Süwer, M.S. Ullah, N. Probul, et al., Privacy-by-design with federated learning will drive future rare disease research, J. Neuromuscul. Dis. 13 (2026) 6–19, https://doi.org/10.1177/22143602241296276.
[72] K.Y.K. So, J.J. Fong, I.P.Y. Lam, et al., Pitfalls during in silico prediction of primer specificity for eDNA surveillance, Ecosphere 11 (2020) e03193, https://doi.org/10.1002/ecs2.3193.
[73] K. Song, H.N. Brochu, Q. Zhang, et al., An in silico analysis of PCR-based monkeypox virus detection assays: a case study for ongoing clinical surveillance, Viruses 15 (2023) 2327, https://doi.org/10.3390/v15122327.
[74] N. Al-Mahrami, S. Al Yazidi, H. Alrashdi, et al., PrimerAST: a predictive machine learning tool for primer design and quality assessment, Sci. Rep. 16 (2026) 14980, https://doi.org/10.1038/s41598-026-38238-8.
[75] R. Andreson, T. Möls, M. Remm, Predicting failure rate of PCR in large genomes, Nucleic Acids Res. 36 (2008) e66, https://doi.org/10.1093/nar/gkn290.
[76] Y. Wei, X. Liu, Y. Mu, et al., From droplets to diagnosis: AI-driven imaging and system integration in digital nucleic acid amplification testing, Biosens. Bioelectron. 288 (2025) 117741, https://doi.org/10.1016/j.bios.2025.117741.
[77] J. Tong, F. Yin, M. Liang, et al., Development of chamber-based digital polymerase chain reaction: from chip to device, TrAC. Trends Anal. Chem. 191 (2025) 118299, https://doi.org/10.1016/j.trac.2025.118299.
[78] S. Wei, S. Wang, Structural stability-aware deep learning: advancing RNA secondary structure prediction, Fourth International Conference on Biomedicine and Bioinformatics Engineering (ICBBE 2024), SPIE, Aug. 2024, pp. 359–369, https://doi.org/10.1117/12.3044258.
[79] R. van Beuningen, K.K. Jim, M. Boot, et al., Development of a large-scale rapid LAMP diagnostic testing platform for pandemic preparedness and outbreak response, Biol. Meth. Protoc. 9 (2024) bpae090, https://doi.org/10.1093/biomethods/bpae090.
[80] C. Conrad, How awareness of a user's cognitive state can improve human-AI teams, in: The Design of Human-Centered Artificial Intelligence for the Workplace, Springer, 2025, pp. 281–291, https://doi.org/10.1007/978-3-031-83512-4_16.
[81] R. Vavekanand, T. Kumar, S. Kumar, et al., Multimodal machine learning approaches in predictive healthcare analytics: a comprehensive survey, Arch. Comput. Meth. Eng. (2026), https://doi.org/10.1007/s11831-026-10560-4.
[82] A. Benani, S. Ohayon, F. Laleye, et al., "Is multimodal better? A systematic review of multimodal versus unimodal machine learning in clinical decision-making (medRxiv), Mar 13 (2025), https://doi.org/10.1101/2025.03.12.25322656.
[83] J. Li, H. Diao, Q. Yu, et al., Knowledge-enhanced explicitly disentangled representation with missing modality for medical image diagnosis, Inf. Fusion. 136 (2026) 104502, https://doi.org/10.1016/j.inffus.2026.104502.
[84] W.F. Faris, N. Gupta, P. Dhadwal, et al., Lightweight AI models for deployment in portable medical devices, in: Proceedings of International Conference on Computational Complexity and Intelligent Algorithms, Springer, 2026, pp. 313–327, https://doi.org/10.1007/978-981-96-7473-2_22.
[85] B.T. Hasan, A.M.A. Al-Sabaawi, Z.J. Al-Araji, Lightweight deep learning for IoT-based medical diagnosis: a survey, Mesop. J. Artif. Intell. Healthc. 2025 (2025) 244–259, https://doi.org/10.58496/mjaih/2025/023.
[86] Y. Wei, S. Luo, C. Xu, et al., SAM-dPCR: accurate and generalist nuclei acid quantification leveraging the zero-shot segment anything model, Adv. Sci. 12 (2025) 2406797, https://doi.org/10.1002/advs.202406797.
[87] A.M. Cotelici, A.T. Bălășoiu, M.V. Boldeanu, et al., Improving the precision of etiological diagnosis in bacterial infections using molecular technologies: a comparative analysis of platforms, AI integration, and point-of-care deployment, Int. J. Mol. Sci. 27 (2026) 4148, https://doi.org/10.3390/ijms27094148.
[88] R. He, D. Hua, W. Wu, et al., Advances of digital detection for foodborne pathogens, Foods 15 (2026) 1250, https://doi.org/10.3390/foods15071250.