The Sydney Journal of Interdisciplinary Engineering

SJIE 1(4) X-X (2025) https://doi.org/XXXX/XXXX

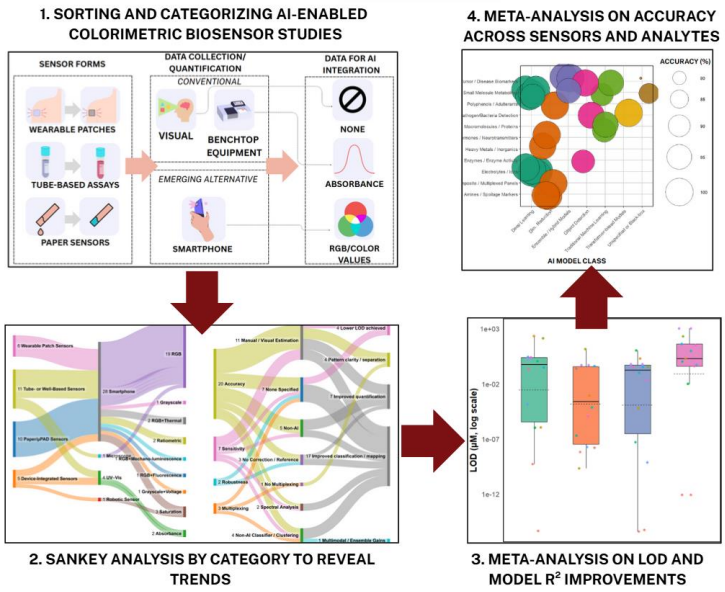
Quantifying the Impact of AI on Colorimetric Biosensor Performance: A Focused Metadata Review

John Salvador Ricacho[1], Gobinath Rajarathnam[1], and Aoni Xu[1]

[1] School of Chemical and Biomolecular Engineering, The University of Sydney

E-mail: jric0361@uni.sydney.edu.au

Graphical Abstract



Abstract

Colorimetric biosensors offer low-cost diagnostics but often suffer from subjective interpretation, environmental variability, and limited quantification. Artificial intelligence (AI) has emerged as a powerful solution, enabling automated analysis of chromogenic outputs captured via smartphones or imaging systems. This meta-analysis reviews 32 studies (2022–2025) applying AI to colorimetric biosensing, comparing performance across model types, sensor formats (e.g., paper, wearable, tube-based), input modalities (e.g., RGB, absorbance), and analyte classes. Key metrics include classification accuracy, regression strength (R^2), and limit of detection (LOD), benchmarked against non-AI and conventional methods.AI-enhanced platforms consistently improved accuracy, with context-specific gains in R^2 and LOD, especially for weak or overlapping signals. Smartphone-based RGB systems dominated but required calibration strategies such as CNN-GRU correction and illumination adjustment. Despite promising results, most studies lacked external validation and relied on supervised learning with small datasets. Semi-supervised approaches and standardized benchmarks are needed to ensure generalizability. Beyond analytical metrics, AI offered faster readouts, automated interpretation, and support for multiplexed sensing. Future directions include integrating augmented reality for enhanced usability and applying AI to sensor design and optimization. Collectively, these advances position AI-enhanced colorimetric biosensors as scalable, field-ready diagnostic tools with growing potential for clinical and environmental deployment

Keywords: colorimetric biosensors, artificial intelligence, machine learning, smartphone diagnostics, point-of-care sensing

1. Introduction

Lack of accurate, acessible, and rapid diagnostics remain a global issue for healthcare especially in remote, resource-constrained settings where over 47% of the global population lacks access to essential diagnostic tools[1]. While conventional laboratory-based diagnostics remian the gold standard, they require sophisticated equipment, trained personnel, and controlled environments. These limitations contribute significantly to delayed diagnoses and diagnostic errors, which are estimated to cause approximately 371,000 deaths and 424,000 permanent disabilities annually in the United States alone[2]. To address these, portable colorimetric biosensors, which are analytical devices that detect presence of target analytes through visible color changes via enzymatic reactions, nanozyme catalysis, or pH-sensitive dyes, have gained prominence as low-cost, easy-to-use alternatives capable of delivering rapid results without the need for laboratory infrastructure[3]. Google Trends data show that global interest in colorimetry more than doubled from late 2021 to early 2025, reflecting growing attention towards visual-based diagnostics[4]. By translating biochemical interactions into observable color changes, they have found applications in diverse settings from at-home glucose monitoring and pregnancy testing to field-based detection of pathogens and heavy metal ions that might be detrimental to health[5]. Moreover, their compatibility with paper-based substrates, lateral flow formats, and nanozyme-enhanced platforms makes them particularly attractive for decentralized healthcare and environmental monitoring[6]. However, despite their significant improvements over traditional diagnostics, colorimetric biosensors face persistent limitations related to subjectivity in optical result interpretation, arising from variations in ambient lighting, camera resolution, user technique, and perceptual bias, which can significantly affect the accuracy and reproducibility of results[7]. This is a hindrance for their widespread adoption in critical clinical or environmental applications where precision and standardisation are essential.

Table 1. Summary of recent reviews on AI-Enabled biosensors and the distinct scope of this work

Year	Focus	Key Insights
2024[7]	AI in biochemical sensors (incl. colorimetric)	Reviewed AI's role across sensing platforms, highlighting accuracy gains and implementation challenges.
2024[3]	AI in electrochemical biosensors	Showed AI improves sensor sensitivity and wearable adaptability.
2024[9]	AI-integrated wound dressings	Reviewed AI-biosensor synergy for wound monitoring and healing prediction.
2023[10]	ML-based sensor arrays for bacterial detection	Surveyed ML-enhanced colorimetric/fluorescent arrays for pathogen classification.
2025 (This review)	AI-enhanced colorimetric biosensors (health & environment)	Conducts first metadata analysis comparing R^2, accuracy, and sensitivity across 30+ studies.

To overcome these challenges, artificial intelligence (AI) has emerged as a transformative solution. By analyzing colorimetric outputs captured via smartphones or imaging devices, AI algorithms provide automated, consistent, and quantitative interpretation of biosensor signals. While previous reviews highlight AI applications in biosensing, few assess its actual performance gains. This review fills that gap through a metadata analysis of recent AI-enhanced colorimetric studies, comparing improvements in sensitivity, accuracy, and regression strength (R^2) over traditional and non-AI methods. Table 1 summarizes prior reviews to contextualize this study's contribution.

2. State-of-the-art of Current Research

This work conducted a metadata analysis of 32 peer-reviewed studies from 2022 to 2025, sourced via Scopus and Google Scholar using combinations of search terms such as "colorimetric biosensor," "AI," "accuracy," and "sensitivity." Studies were included if they employed artificial intelligence (machine learning or deep learning) for the interpretation of colorimetric biosensor outputs and reported at least one quantitative performance metric (e.g., accuracy, sensitivity, or R^2). Data were manually extracted on sensor type, sample source, analyte, AI model, and comparative improvement over non-AI or traditional methods. A comprehensive table detailing these 32 studies is shown in Table S1 (supplementary).

2.1. Sensor Architecture – Form Factor, Platform, and AI Data Utilization

Figure 2.4.a shows sensor architectures across the 32 studies prioritized cost-effectiveness, portability, and user-friendliness, which are qualities best demonstrated by paper-based sensors (10 studies), wearable microfluidic patches (6), and tube/well-based formats (11), collectively accounting for over 85% of sensor form factors. In comparison to conventional laboratory-based diagnostics, these form factors drastically reduce overheads in terms of materials and logistics, enabling decentralized testing. Device-integrated sensors (5 studies), while offering superior performance via embedded optics or processors, still lack scalability due to their high cost and need for specialized maintenance. Smartphones were overwhelmingly used for signal collection (28 of 32 studies), outpacing other platforms like scanners (4), robotic sensors (1), and microscopes (1), due to their widespread accessibility, built-in cameras, and ability to process or upload images in real time. This sensor architecture across these studies is visualized in Figure 2.1.

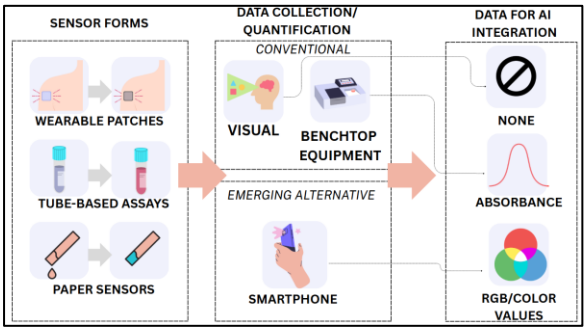


Figure 2.1. Architecture of AI-enabled colorimetric sensors

RGB was the main input for AI models (19 studies), followed by grayscale (2), absorbance (3), and multimodal setups like RGB with thermal, mechanoluminescence, or fluorescence (1 each). Its appeal lies in smartphone compatibility and suitability for CNNs that process spatial and color features. Absorbance-based methods are more robust but rely on non-portable, specialized tools. RGB's sensitivity to lighting and device variability reduces reliability without normalization, used in only a few studies (some in Table 2.1).

These steps are key to improving consistency in real-world settings. Overall, the move toward RGB-smartphone-AI systems supports scalable diagnostics, but stronger standardization is still needed to match lab-grade performance.

Table 2.1 Color correction strategies applied across selected studies

Study	Color correction applied
Wang et al.[11]	Trained a CNN-GRU model to adjust for ambient light and pH variation
Ghateii and Jahanshashi [36]	Used flash/no-flash subtraction and lab color space conversion to stabilize lighting conditions
Liu et al[42]	Applied pixel-wise color correction using a 24-color checker to calibrate camera-based inputs

2.2 Purpose and Sample Type – Monitoring Targets, Matrices, and Analytes

Figure 2.4.b reveals that sensors are mainly applied to clinical diagnostics (10/32 studies), metabolic monitoring (7), and food safety (6), with fewer targeting pathogens (4), cellular assays (2), or multiplex panels (1). This mirrors the prevalence of accessible samples like urine (4), sweat (3), saliva, and tears, ideal for wearable or point-of-care use. However, this also suggests an application bias, favoring well-characterized analytes in controlled settings. Food and environmental samples (9 studies combined), which present greater matrix complexity and signal noise, remain underrepresented despite being where AI's disambiguation strengths are most needed. Current trends favor feasibility over impact, applying AI where outputs are already interpretable rather than where its value is most critical.

Notably, many AI models have been applied to analytes that already produce vivid and monotonic color changes, such as glucose and pH, where human-readable output is already largely feasible. While this enables automation and precision, it may underutilize AI's potential. As shown in Figure 2.2, analytes like HDL, LDL, and troponin exhibit weaker or grayscale transitions that are far less distinguishable visually. These cases present the strongest justification for AI integration yet remain underrepresented. Rather than reinforcing already discernible signals, AI's role should be expanded to support analytes with ambiguous visual responses, where its capacity for pattern recognition and subtle gradient differentiation can meaningfully extend the reach of colorimetric sensing.

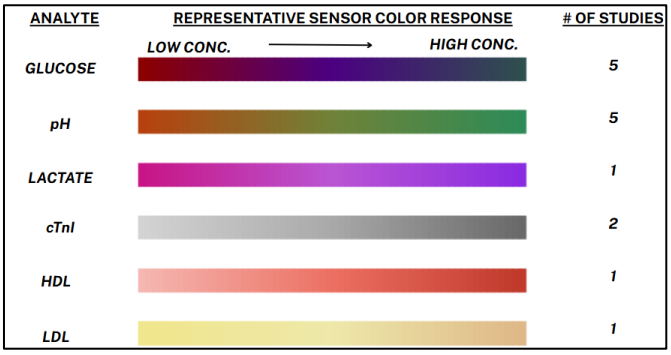


Figure 2.2. Colorimetric responses for selected analytes.

2.3 AI Use Case and Model – Task Types and Algorithms Employed

AI in colorimetric biosensing has mainly focused on regression (18 of 32 studies) and classification (13), aligning with the direct relationship between color change and either concentration or categorical outcome. Regression typically maps RGB patterns to analyte levels, while classification supports test result interpretation. These applications suit sensors targeting analytes with clear, monotonic color shifts like glucose or pH. However, this also reflects a cautious approach where AI is often applied where signal-response relationships are already well defined. More advanced tasks like clustering, anomaly detection, or multimodal fusion remain rare, despite their potential for handling complex or noisy signals.

Figure 2.3 shows a mismatch between AI task complexity and the models used in reviewed studies. Simpler regression tasks were most common and often addressed with traditional ML models like random forests, even when signals were nonlinear or noisy. Deep learning was more common in classification tasks, particularly for spatial data, but rarely used for complex tasks like object detection or multimodal fusion. For example, Yu et al.[27] used an ANN for RGB-thermal fusion but didn't apply advanced architectures like attention or transformers. Unsupervised methods like PCA or t-SNE were limited to visualization. This suggests model selection is often based on familiarity, not task fit. As a result, underspecified models may limit performance in complex or noisy settings and reduce generalizability outside the lab. Treating model architecture as a key design element, aligned with task demands and supported by benchmarking, will be essential for advancing AI in biosensing.

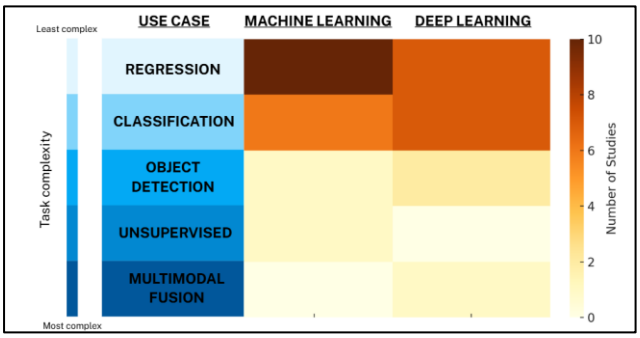


Figure 2.3. AI model use by task type, showing ML dominates regression while DL is underused in complex tasks.

2.4 Performance Improvement – Gains Attributed to AI and Benchmarks

Across all 32 studies, AI integration was credited with enhancing sensor performance across multiple axes. The most reported gains were improved accuracy (~20 studies), enhanced sensitivity or lower limits of detection (~7 studies), faster or automated interpretation (~3 studies), and improved pattern resolution for multiplexed or overlapping signals (~4 studies). AI enabled detection of subtle analyte differences, automated endpoint interpretation, and separation of overlapping outputs in multi-analyte sensors. While about 7 studies lacked a baseline comparison, those that did consistently showed AI outperforming visual reads, thresholds, or uncorrected data. Table 2.2 highlights four representative examples. Cui et al.[12] used YOLOv5 to improve bacterial classification to 95%.

Yu et al.[27] combined colorimetric and thermal signals via ANN to surpass LOD for cardiac troponin. In Zheng et al's work[16], CNNs reduced assay readout time withing minutes to seconds, while originally taking hours. Ranbir et al[25]. and Singh et al.[30] used PCA-LDA to fully separate volatile amines in meat, showing AI's strength in multiplex detection. These examples illustrate both performance gains and how targeted AI use can expand the utility of colorimetric sensors in real-world settings. However, while showing these gains, a more quantitative approach is required to fully grasp the importance of AI in colorimetric biosensing, as explored in subsequent section.

2.4 Meta-analysis of performance improvements

2.4.1. Limit of detection (LOD)

We compared LOD values across studies, as LOD reflects the lowest detectable concentration above background noise and is key to assessing sensor sensitivity. This helps determine whether AI meaningfully improves detection limits in real-world use.

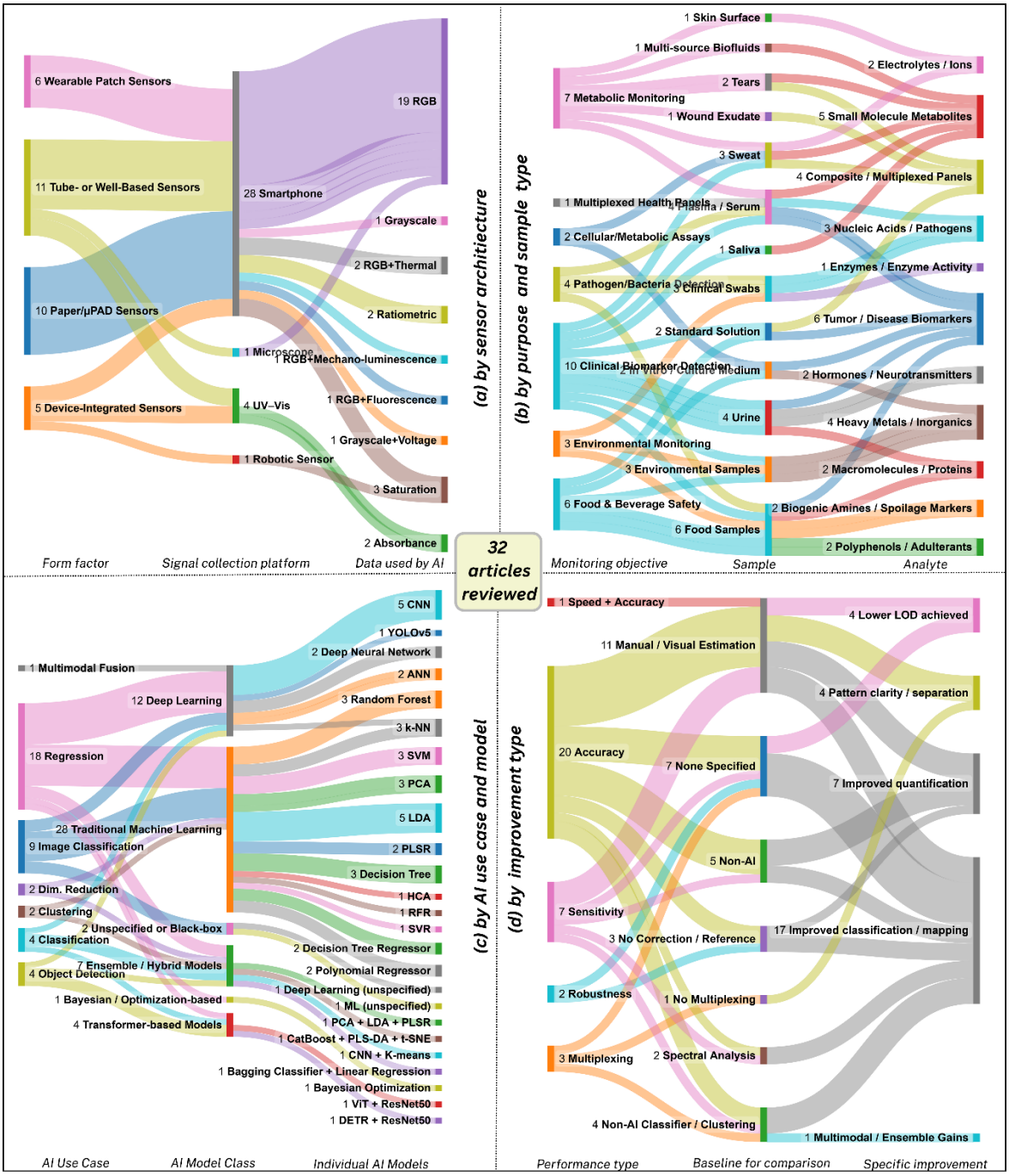


Figure 2.4 Sankey diagram for AI-enhanced colorimetric biosensor studies sorted by (a) sensor architecture; (b) purpose and sample types, (c) AI use cases, model classes, and algorithms; and (d) performance improvements (comprehensive table in Table S1)

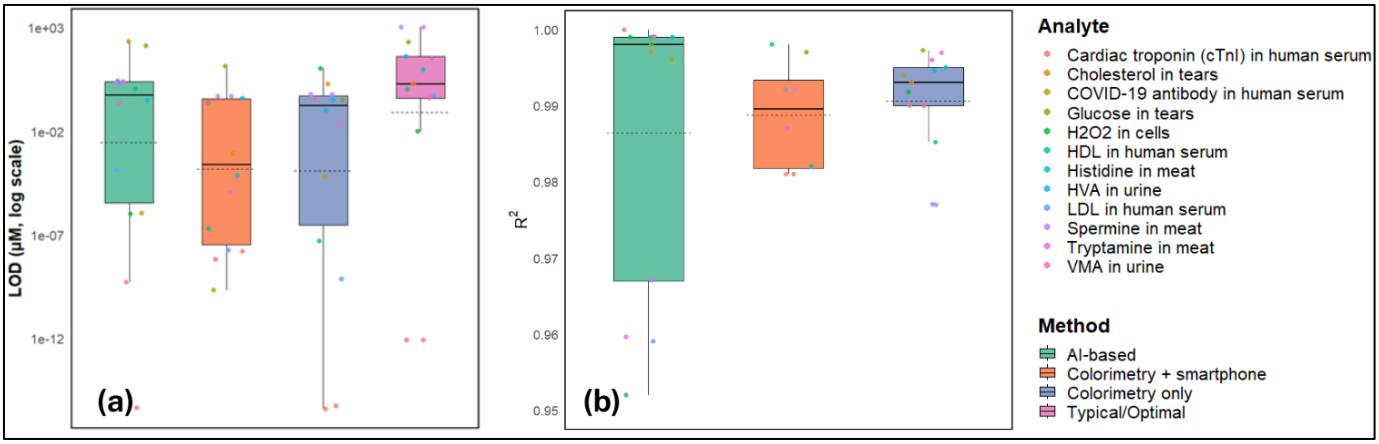


Figure 2.5. Log-scale distribution of LOD values across colorimetric detection methods, highlighting analyte-level spread and comparison to typical/optimal reference values

LOD data from AI-enhanced studies were grouped into four categories: AI-based, colorimetry only, colorimetry plus smartphone (non-AI), and typical values in human or food samples. Each analyte within a study was treated as a separate data point. For baseline methods lacking internal controls, LODs were sourced from recent reviews or similar studies. Full data appear in Supplementary Table S1.

As shown in Figure 2.5.a, AI-based platforms had slightly lower median LODs than typical concentrations in human and health samples, highlighting significance in diagnostics, but variability was high across all groups, especially compared to non-AI smartphone-assisted methods. Mann–Whitney U tests (Table S4) showed no statistically significant differences (all $p > 0.15$), indicating that AI alone doesn't consistently improve sensitivity. Outliers like del Real Mata et al's[13] 1 pM H_2O_2 detection with a plasmonic sensor and random forest model, or Yu et al's[27] 10.8 pg/mL troponin detection using ANN fusion, highlight AI's potential under optimized setups. However, factors like sensor materials, analyte properties, and sample matrices often have greater influence. AI was most impactful in cases with overlapping or faint color signals, e.g. Cui et al's use of YOLOv5 for low-level bacterial HAase, and Ranbir et al's[25.] and Singh et al's.[30] PCA-LDA models resolving mixed biogenic amines. In contrast, analytes with strong color change like glucose or pH showed minimal LOD gains, though AI improved consistency and automation. These results suggest AI should be applied selectively, especially for low-contrast or nonlinear signals. Broader adoption will require better benchmarking, task-specific AI design, real-world validation, and comparison to regulatory standards or reference methods.

2.4.2. Model R^2 values

We included R^2 comparisons across studies as it reflects how well a sensor's output follows analyte concentration trends making it an essential indicator of dose-response consistency, even if not a direct accuracy measure. R^2 data were grouped by method type (AI-based, colorimetry-only, and smartphone-assisted non-AI) and are summarized in Table S3 and visualized in Figure 2.5.b. AI-based platforms showed higher average R^2 values (0.952–0.9999) and wider spread than conventional methods, with several achieving near-perfect calibration under controlled conditions. However, Mann–Whitney U tests (Table S5) indicated these differences weren't statistically significant ($p = 0.075$ vs. smartphone; $p = 0.14$ vs. colorimetry-only), suggesting AI doesn't consistently improve regression fit across all cases. The best R^2 values were seen in studies with controlled imaging, high signal-to-noise ratios, or carefully curated datasets. Study 18 reached $R^2 = 0.9999$ for cardiac troponin I using an ANN with thermal and color fusion, while Study 13 achieved robust fits for glucose and cholesterol using ensemble models that corrected ambient lighting. By contrast, non-AI methods, especially smartphone-only approaches, showed greater performance drops under uncontrolled conditions, with R^2 values around 0.79–0.80 for LDL and HDL, likely due to lighting variability. These findings suggest AI's greatest strength lies in stabilizing regression under noisy or nonlinear signal conditions. However, high R^2 alone is not sufficient. Some colorimetry-only systems still performed well for monotonic, high-contrast analytes, highlighting the continued importance of sensor chemistry. To ensure robust performance, future work should combine R^2 with broader metrics like residual analysis, external validation, and real-world testing. Overreliance on R^2 may inflate confidence, particularly in the absence of clinical or field verification.

2.4.2. Accuracy

Unlike limit of detection (LOD) and regression metrics such as R^2, classification accuracy lacks a consistent universal baseline in biosensing literature. The diversity of decision thresholds, analyte classes, and labeling protocols across studies means that accuracy figures are highly context dependent. As such, we do not compare absolute values across platforms. Instead, we focus on within-dataset patterns observed across sensor architectures, analyte groups, and AI model classes, as summarized in Figure 2.6a and Supplementary Table S6. Overall, AI-enhanced biosensors consistently outperformed non-AI platforms, with the majority of AI-based systems achieving accuracies above 90%, and several reaching the 100% benchmark across diverse sensing contexts. These included both deep learning and hybrid ensemble methods, suggesting the benefits of nonlinear pattern recognition, especially when signal variability or interference is present.

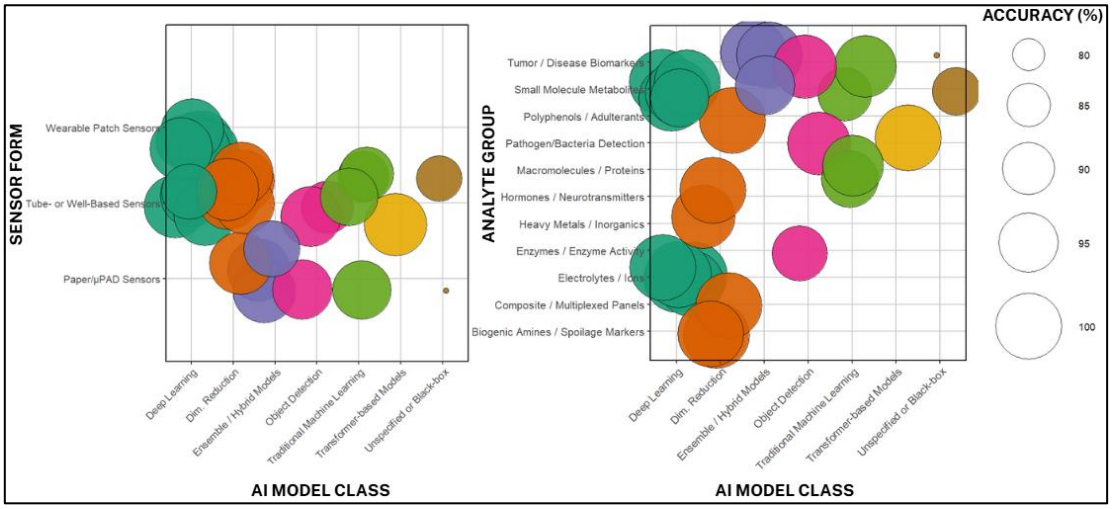


Figure 2.6. Bubble plot of classification/quantification accuracy of AI-enhanced colorimetric biosensors, mapped across sensor form factors (left) and analyte groups (right) by AI model class

At the architecture level, wearable patch sensors, when paired with CNN-based models, demonstrated high robustness and accuracy, often exceeding 95% for multi-biomarker sweat patches. Study 14 achieved 100% classification for glucose, pH, and lactate, enabled by a VGG16 CNN that captured subtle differences in spatial signal distributions under ambient conditions. Similarly, paper/µPAD sensors paired with traditional ML models (e.g., Random Forest, SVM) also performed well, particularly for urinary and metabolic analytes, where structured chromogenic arrays generated reproducible color fingerprints. Study 11, for example, achieved 97% accuracy in urinary tract infection classification using an SVM-RF ensemble.Among analyte categories, tumor and cardiac biomarkers benefited most from AI integration. The fusion of thermal and optical signals in Study 18, using an ANN, yielded accurate discrimination of cardiac troponin I (cTnI), reinforcing the strength of multimodal biosensing for critical clinical analytes . Additionally, for biogenic amines, LDA-based models maintained >95% accuracy, even under food matrix variability.

In contrast, non-AI systems, especially those relying on smartphone cameras with simple thresholding or raw RGB interpretation, showed greater susceptibility to lighting inconsistencies, with accuracy often falling to the 85–90% range. These limitations were particularly evident in complex backgrounds like food spoilage detection or overlapping chromophores, where AI methods (e.g., PCA-LDA fusion) restored classification clarity. From the sensor form perspective, tube- or well-based formats showed generally stable accuracy due to controlled optics, though wearable and paper-based formats 6abelled them when enhanced by AI. Notably, the highest accuracies clustered in deep learning and object detection classes (Figure 2.6a left panel), reflecting their superior ability to extract spatial and contextual features from raw image data. Together, these trends suggest that while chemical design and sensor chemistry remain foundational, AI integration—especially through CNNs, hybrid models, and transformer-based architectures—can significantly amplify diagnostic reliability, especially under variable environmental or user-handling conditions. Future work should explore adaptive learning for personalized calibration and establish standardized accuracy benchmarks across sensor classes.

3. Synthesis and outlook

1. On the use of smartphones and calibration needs- The collected studies make clear that coupling AI with colorimetric biosensors can dramatically enhance their capabilities, turning simple color changes into rich quantitative and actionable data. A unifying theme is the leveraging of ubiquitous hardware, particularly smartphones , as both the data acquisition device and computation platform. This convergence, seen in roughly 90% of the articles, underscores a practical advantage: AI algorithms deployed on consumer smartphones can transform point-of-care diagnostics, allowing immediate analysis in the field. However, this shift toward RGB smartphone-based pipelines brings new challenges in data normalization. Different phone cameras and ambient lighting conditions can skew color readings, requiring robust calibration to ensure reproducibility[67]. Encouragingly, several teams have introduced clever calibration techniques to tackle this issue. For example, cloud-connected analysis frameworks now incorporate hybrid models (CNNs coupled with recurrent networks) to auto-correct for illumination variances and sensor-specific biases. Such approaches (e.g. a multichannel CNN-GRU pipeline) have achieved R^2 values ~0.99 by learning to adjust for color temperature differences in images, effectively standardizing results across varying conditions. Moving forward, continued innovation in on-device calibration (from one-time color card references to real-time algorithmic corrections) will be essential to fully capitalize on smartphone-enabled AI sensing.

2. On generalizability and data-efficient modelling- Despite the impressive performance gains reported, most studies lack rigorous external validation, highlighting a critical gap between controlled experiments and real-world deployment. Typically, models are trained and tested on the same lab-generated dataset; few works verify that an AI model trained on one device or sample set holds up on others. This absence of external validation and cross-platform testing raises concerns about generalizability, an issue that future research must address by incorporating independent test sets, multi-center trials, or reference sample exchanges. Likewise, the underuse of semi-supervised learning and data augmentation is notable. Many AI models for colorimetric sensing rely on relatively small labeled

datasets, yet few studies leverage the abundance of unlabeled data or synthetic data generation to improve model robustness. Introducing semi-supervised algorithms (which can learn from unlabeled color images) or augmentation techniques (to simulate variations in hue, intensity, backgrounds, etc.) could significantly enhance model resilience to real-world variability at minimal cost. Another insight from our meta-analysis is that AI's added value appears tied less to the analyte type and more to the ambiguity of the signal. In other words, when an assay produces straightforward, high-contrast color changes (e.g. a single intense color shift for a positive result), traditional analysis may suffice. But as the color outputs become more complex, such as subtle gradations, multi-analyte sensor arrays, or overlapping chromatic responses, advanced machine learning yields disproportionate benefits[68]. Indeed, deep learning models excel at deciphering high-dimensional color patterns that humans or simple algorithms struggle to interpret. This trend suggests that future developers should strategically deploy AI in scenarios of inherent signal complexity or uncertainty, where its pattern-recognition strengths are most impactful. It also implies that reporting performance as a function of assay complexity (rather than only by analyte category) could be a more meaningful way to evaluate new AI-enhanced biosensors.

3. On practical gains: speed, multiplexing, and robustness- From a practical standpoint, AI-driven colorimetric analysis offers improvements that extend beyond raw analytical metrics, contributing to better usability and reliability of biosensors. One clear advantage is speed: once trained, an AI model can interpret a sensor's color output in milliseconds, potentially enabling near real-time readouts and quicker decision-making in point-of-care settings. In some cases, algorithms can even detect partial color changes before a reaction is fully complete, shortening the time-to-result. Another benefit is the capacity for multiplexed detection, that is, analyzing multiple indicators simultaneously. Traditional colorimetric assays struggle when multiple test spots or mixed-color outputs must be interpreted at once, whereas machine learning can untangle such composite signals with high accuracy. For example, neural network models have distinguished multiple antibody responses in a single assay with ~89% accuracy, outperforming conventional methods by a significant margin[68]. In general, as more analytes are encoded into color-based tests, AI will be instrumental in accurately classifying outcomes across a multidimensional color space. Equally important is the robustness that AI brings: sophisticated models can accommodate variability in sample quality or environmental conditions (such as inconsistent lighting or user handling) better than rigid threshold-based interpretations. Notably, convolutional neural networks have maintained strong performance even when images are noisy or under suboptimal lighting, a resilience crucial for real-world applications. This robustness reduces the incidence of false negatives or false positives caused by minor perturbations, thus improving trust in home or field deployments.

4. On the horizon: integration with AR/VR and digital design- Looking towards the horizon, there are exciting opportunities to integrate emerging technologies like augmented and virtual reality (AR/VR) with AI-based colorometric sensing. Early demonstrations have shown that AR smartphone apps can overlay interpretive guidance or even embed fiducial markers into the test to aid real-time result reading. In the future, a user might simply point a phone at a paper sensor and see a quantified result or risk assessment pop up instantly via AR, lowering the barrier to accurate self-testing. VR environments could also serve as training tools, simulating a wide range of colorimetric outcomes for clinicians or as a platform to virtually prototype sensor designs. Moreover, AI itself can be applied beyond analysis – for instance, using machine-learning optimization to design better colorimetric assays (selecting optimal reagent combinations or layout to maximize signal differentiation) or to create digital twins that predict how a sensor will behave under various scenarios. These exploratory directions, while in nascent stages, underscore the expansive potential at the interface of smart algorithms and biosensing. In summary, the future outlook for AI-enhanced colorimetric biosensors is one of continued convergence by merging accessible hardware, powerful algorithms, and user-centric innovations to deliver faster, multiplexed, and more robust diagnostic solutions. The next few years will likely witness not only incremental performance improvements but also a maturing of the field through standardized evaluation protocols, open datasets for model training, and perhaps the advent of intelligent sensors that learn and adapt during use.

4. Acknowledgments

The authors acknowledge the use of artificial intelligence tools for language refinement and copyediting during manuscript preparation. All content and interpretations remain the sole responsibility of the authors.

References

1. Fleming, K. A. et al. The Lancet Commission on diagnostics: transforming access to diagnostics. *Lancet* **2021**, *398*(10315), 1997–2050. https://doi.org/10.1016/S0140-6736(21)00673-5

2. Newman-Toker, D. E. et al. Serious misdiagnosis-related harms in malpractice claims: The "Big Three" – vascular events, infections, and cancers. *BMJ Qual. Saf.* **2023**, *32*, 100–112.

3. Sun, Z., Zhang, B., Tu, H., Pan, C., Chai, Y., & Chen, W. Advances in colorimetric biosensors of exosomes: novel approaches based on natural enzymes and nanozymes. *Nanoscale* **2024**, *16*, 1005–

4. Google Trends. *Colorimetry – Worldwide, 2021–2025*. https://trends.google.com (accessed 4 May 2025).

5. Yang, F.-Q., & Ge, L. Colorimetric Sensors: Methods and Applications. *Sensors* **2023**, *23*(5), 2749

6. Cai, Y., Wang, Y., Wu, Y., et al. Current trends in colorimetric biosensors using nanozymes for detecting biotoxins. *Analyst* **2024**, *149*, 1234–1245.

7. Flynn, C. D. & Chang, D. Artificial Intelligence in Point-of-Care Biosensing: Challenges and Opportunities. *Diagnostics* **14**, 1100 (2024).

8. Nur, S., Salleh, M., Rozan Mohamad Yunus & Halimah Badioze Zaman. Artificial Intelligence−Powered Electrochemical Sensor: Recent Advances, Challenges, and Prospects. *Heliyon* **10**, e37964–e37964 (2024).

9. Prakashan, D., Kaushik, A. & Gandhi, S. Smart sensors and wound dressings: Artificial intelligence-supported chronic skin monitoring – A review. Chemical Engineering Journal 497, 154371 (2024).

10. Yang, C. & Zhang, H. A review on machine learning–powered fluorescent and colorimetric sensor arrays for bacteria identification. Microchimica Acta 190, (2023).

11. Wang, Z. *et al.* An artificial intelligence-assisted microfluidic colorimetric wearable sensor system for monitoring of key tear biomarkers. *Npj Flexible Electronics* **8**, (2024).

12.Cui, R. *et al.* AI-assisted smartphone-based colorimetric biosensor for visualized, rapid and sensitive detection of pathogenic bacteria. *Biosensors and Bioelectronics* **259**, 116369 (2024).

13. del Real Mata, C. *et al.* AI-Assisted Plasmonic Enhanced Colorimetric Fluidic Device for Hydrogen Peroxide Detection from Cancer Cells. *Advanced Materials Technologies* **10**, (2024).

14.Sun, S. *et al.* BioCosMe: Lip-based Cosmetics with Colorimetric Biosensors for Salivary Analysis using Deep Learning. *Proceedings of the 2024 ACM International Symposium on Wearable Computers* 32–39 (2024) doi:https://doi.org/10.1145/3675095.3676610.

15.Hassani-Marand, M., Jafarinejad, S. & Hormozi-Nezhad, M. R. An AI-enabled multi colorimetric sensor array: Towards rapid and noninvasive detection of neuroblastoma urinary markers. *Sensors and Actuators B: Chemical* **396**, 134571 (2023).

16. Xin Ting Zheng *et al.* Battery-free and AI-enabled multiplexed sensor patches for wound monitoring. *Battery-free and AI-enabled multiplexed sensor patches for wound monitoring* **9**, (2023).

17. Tong, H. *et al.* Artificial intelligence-assisted colorimetric lateral flow immunoassay for sensitive and quantitative detection of COVID-19 neutralizing antibody. *Biosensors and Bioelectronics* **213**, 114449 (2022).

18. Wansadaj Jaroenram *et al.* One-step colorimetric isothermal detection of COVID-19 with AI-assisted automated result analysis: A platform model for future emerging point-of-care RNA/DNA disease diagnosis. **249**, 123375–123375 (2022).

19. Siew, J., T. Malathi Thevarajah, Chang, S. W. & Khor, S. M. Paper-based multiplexed colorimetric biosensing of cardiac and lipid biomarkers integrated with machine learning for accurate acute myocardial infarction early diagnosis and prognosis. *Sensors and Actuators B Chemical* **394**, 134403–134403 (2023).

20.Yang, J. *et al.* Machine Learning-Assistant Colorimetric Sensor Arrays for Intelligent and Rapid Diagnosis of Urinary Tract Infection. *ACS Sensors* **9**, 1945–1956 (2024).

21.Feng, F. *et al.* Artificial intelligence-assisted colorimetry for urine glucose detection towards enhanced sensitivity, accuracy, resolution, and anti-illuminating capability. *Nano Research* (2023) doi:https://doi.org/10.1007/s12274-022-5311-5.

22. Baştürk, M., Yüzer, E., şen, M. & Kılıç, V. Smartphone-Embedded Artificial Intelligence-Based Regression for Colorimetric Quantification of Multiple Analytes with a Microfluidic Paper-Based Analytical Device in Synthetic Tears. *Advanced Intelligent Systems* (2024) doi:https://doi.org/10.1002/aisy.202400202.

23.Liu, Z. *et al.* Explainable Deep-Learning-Assisted Sweat Assessment via a Programmable Colorimetric Chip. *Analytical Chemistry* **94**, 15864–15872 (2022).

24. Do Hyeon Jung *et al.* Automatic quantification of living cells via a non-invasive achromatic colorimetric sensor through machine learning-assisted image analysis using a smartphone. *Chemical engineering journal* **450**, 138281–138281 (2022).

25.Ranbir, Singh, G., Singh, H., Kaur, N. & Singh, N. Azodye-based colorimetric sensor array for identification of biogenic amines: Food forensics by portable RGB-based signal readout. *Sensors and Actuators B: Chemical* **387**, 133794 (2023).

26. Li, H. *et al.* Porphyrin and pH sensitive dye-based colorimetric sensor array coupled chemometrics for dynamic monitoring of tea quality during ultrasound-assisted fermentation. *Microchemical Journal* **197**, 109813 (2023).

27. Yu, Z., Gong, H., Li, M. & Tang, D. Hollow prussian blue nanozyme-richened liposome for artificial neural network-assisted multimodal colorimetric-photothermal immunoassay on smartphone. *Biosensors and Bioelectronics* **218**, 114751 (2022).

28. Zhang, J. *et al.* Explainable Deep Learning-Assisted Self-Calibrating Colorimetric Patches for In Situ Sweat Analysis. *Analytical Chemistry* **96**, 1205–1213 (2024).

29. Zhu, X., Li, T., Hai, X. & Bi, S. A nanozyme-based colorimetric sensor array as electronic tongue for thiols discrimination and disease identification. *Biosensors and Bioelectronics* **213**, 114438 (2022).

30. Singh, H., Singh, G., Kaur, N. & Singh, N. Pattern-based colorimetric sensor array to monitor food spoilage using automated high-throughput analysis. *Biosensors and Bioelectronics* **196**, 113687 (2022).

31. Feng, H. *et al.* Bimodal Visual Sensors Based on Mechanoluminescence and Biosensing for Artificial Intelligence-Assisted Orthodontics. *Advanced Functional Materials* (2024) doi:https://doi.org/10.1002/adfm.202416437.

32. Zhai, C. *et al.* An enzyme response-regulated colorimetric assay for pattern recognition sensing application using biomimetic inorganic-protein hybrid nanoflowers. *Chemical Engineering Journal* **431**, 134107 (2022).

33. Yang, X. *et al.* A colorimetric sensor array based on peroxidase activity nanozyme for the highly efficient differential sensing of tea polyphenols and Tieguanyin adulteration. *Food Chemistry* **432**, 137265 (2023).

34. Kong, Y. *et al.* Artificial neural network-facilitated V2C MNs-based colorimetric/fluorescence dual-channel biosensor for highly sensitive detection of AFB1 in peanut. *Talanta* **266**, 125056 (2024).

35. Chen, Y. *et al.* Robot-accelerated development of a colorimetric CO2 sensing array with wide ranges and high sensitivity via multi-target Bayesian optimizations. *Sensors and Actuators B: Chemical* **390**, 133942 (2023).

36. Solaleh Ghateii & Amir Jahanshahi. Colorimetric detection of glucose with smartphone-coupled µPADs: harnessing machine learning algorithms in variable lighting environments. *Sensors and actuators. B, Chemical* **400**, 134835–134835 (2024).

37. Hassani-Marand, M., Fahimi-Kashani, N. & Hormozi Nezhad, M. R. Machine-Learning Assisted Multiplex Detection of Catecholamine Neurotransmitters with a Colorimetric Sensor Array. *Analytical Methods* (2023) doi:https://doi.org/10.1039/d2ay01797k.

38. Lee, T. *et al.* A regression-based machine learning approach for pH and glucose detection with redox-sensitive colorimetric paper sensors. *Analytical Methods* **14**, 4749–4755 (2022).

39. Xu, J. *et al.* Machine Learning-Assisted Portable Dual-Readout Biosensor for Visual Detection of Milk Allergen. *Nano Letters* (2025) doi:https://doi.org/10.1021/acs.nanolett.5c01001.

40. Bhatt, S., Kumar, S., Gupta, M. K., Datta, S. K. & Dubey, S. K. Colorimetry-based and smartphone-assisted machine-learning model for quantification of urinary albumin. *Measurement Science and Technology* **35**, 015030–015030 (2023).

41. Mizaj Shabil Sha *et al.* A smartphone-interfaced, low-cost colorimetry biosensor for selective detection of bronchiectasis*via*an artificial neural network. *RSC advances* **12**, 23946–23955 (2022).

42. Liu, W. *et al.* Machine-Learning-Based Colorimetric Sensor on Smartphone for Salivary Uric Acid Detection. *Ieee.org* (2023) doi:https://doi.org/10.21227/8ad0-mw22.

43.Xi, X. *et al.* Preparation of Au/Pt/Ti3C2Cl2 nanoflakes with self-reducing method for colorimetric detection of glutathione and intracellular sensing of hydrogen peroxide. *Carbon* **197**, 476–484 (2022).

44.Peng, Z. *et al.* Rapid colorimetric detection of H2O2 in living cells and its upstream series of molecules based on oxidase-like activity of CoMnO3 nanofibers. *Sensors and Actuators B: Chemical* **382**, 133540 (2023).

45.Kasatkin, A. & Urakov, A. Effect of hydrogen peroxide on erythrocyte temperature in vitro. *Chemico-Biological Interactions* **354**, 109837 (2022).

46.Miri, A., Orouji, A. & Hormozi-Nezhad, M. R. Etched-suppressed gold nanorods providing highly distinctive plasmonic patterns: Towards multiplex analysis of neuroblastoma biomarkers. *Analytica Chimica Acta* **1325**, 343119 (2024).

47.Surasak Kasetsirikul *et al.* Detection of the SARS-CoV-2 humanized antibody with paper-based ELISA. *Analyst* **145**, 7680–7686 (2020).

48.Ulloa-Gomez, A. M., Agredo, A., Lucas, A., Somvanshi, S. B. & Stanciu, L. Smartphone-based colorimetric detection of cardiac troponin T via label-free aptasensing. *Biosensors and Bioelectronics* **222**, 114938 (2022).

49. University of Rochester Medical Center. *Rochester.edu* https://www.urmc.rochester.edu/encyclopedia/content?contenttypeid=167&contentid=troponin (2024).

50. Li, G. *et al.* Colorimetric aptasensors for sensitive low-density lipoprotein detection based on reduced oxide graphene@molybdenum disulfide-ferrocene nanosheets with peroxidase-like activity. *Analytical Methods* (2024) doi:https://doi.org/10.1039/d4ay01648c.

51.Pakira, V., Agarwal, R., Chatterjee, S., Mukherjee, A. & Chakraborty, S. Lipidest: a lipid profile screening test under extreme point of care settings using a portable spinning disc and an office scanner. *Analytical Methods* **15**, 2427–2440 (2023).

52.Johns Hopkins Medicine. Lipid panel. *John Hopkins Medicine* https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/lipid-panel (2024).

53.Chunta, S., Suedee, R. & Lieberzeit, P. A. High-density lipoprotein sensor based on molecularly imprinted polymer. *Analytical and Bioanalytical Chemistry* **410**, 875–883 (2017).

54. Fan, J., Zhou, K. & Wang, J. Glucose Oxidase Coupling with Pistol-Like DNAzyme Based Colorimetric Assay for Sensitive Glucose Detection in Tears and Saliva. *Applied Biochemistry and Biotechnology* (2024) doi:https://doi.org/10.1007/s12010-024-05046-7.

55. Kap, Ö., Kılıç, V., Hardy, J. G. & Horzum, N. Smartphone-based colorimetric detection systems for glucose monitoring in the diagnosis and management of diabetes. *Analyst* **146**, 2784–2806 (2021).

56. Sen, D. K. & Sarin, G. S. Tear glucose levels in normal people and in diabetic patients. *British Journal of Ophthalmology* **64**, 693–695 (1980).

57. Alle, M. *et al.* Gold nanoparticles spontaneously grown on cellulose nanofibrils as a reusable nanozyme for colorimetric detection of cholesterol in human serum. *International Journal of Biological Macromolecules* **201**, 686–697 (2022).

58. Shi, Z. *et al.* Highly sensitive and selective colorimetric and fluorescent dual-readout assay for the analysis of spermine, spermidine and cysteine. *Dyes and Pigments* **220**, 111659 (2023).

59. Wei, Y. *et al.* Smartphone-integrated colorimetric sensor for rapid and highly selective detection of spermine in food based on the laccase-mimicking activity of flower-shaped Mn3O4 nanoparticles. *Microchemical Journal* **198**, 110148–110148 (2024).

60.Naila, A., Flint, S., Fletcher, G., Bremer, P. & Meerdink, G. Control of Biogenic Amines in Food-Existing and Emerging Approaches. *Journal of Food Science* **75**, R139–R150 (2010).

61. Fan, M. *et al.* Highly sensitive detection of tryptophan (Trp) in serum based on diazo-reaction coupling with Surface-Enhanced Raman Scattering and colorimetric assay. *Analytica Chimica Acta* **1119**, 52–59 (2020).

62. Xie, C. *et al.* Fast detection of tryptamine in meat products with azide-functionalized covalent organic frameworks confined in molecularly imprinted polymers. *Food Chemistry* **452**, 139527 (2024).

63.Schirone, M. *et al.* Biogenic Amines in Meat and Meat Products: A Review of the Science and Future Perspectives. *Foods* **11**, 788 (2022).

64. Saadati, A. *et al.* Optical discrimination of histamine and ethylenediamine in meat samples using a colorimetric affordable test strip (CATS): introducing a novel lab-on paper sensing strategy for low-cost ensuring food safety by rapid and accurate monitoring of biogenic amines. *RSC Advances* **14**, 8602–8614 (2024).

65.Wang, J. *et al.* DNAzyme-based and smartphone-assisted colorimetric biosensor for ultrasensitive and highly selective detection of histamine in meats. *Food Chemistry* **435**, 137526 (2024).

66.Michalski, M., Pawul-Gruba, M. & Madejska, A. Histamine contents in raw long-ripening meat products commercially available in Poland. *Journal of Veterinary Research* **65**, 477–481 (2021).

67. Russell, S. M., Doménech-Sánchez, A. & Roberto. Augmented Reality for Real-Time Detection and Interpretation of Colorimetric Signals Generated by Paper-Based Biosensors. *ACS Sensors* **2**, 848–853 (2017).

68.Han, G.-R. *et al.* Machine learning in point-of-care testing: innovations, challenges, and opportunities. *Nature Communications* **16**, (2025).

Table S1. Comprehensive table for the 32 reviewed studies on AI-enabled colorimetric biosensor

Sensor Type	Analyte	Sample Source	Detection Mechanism	Signal collection platform for AI	AI Model	Role of AI (with Subgroup Tags)	Reported LOD	Classification Accuracy	Regression Accuracy (inverted MAE or similar)	Fit Quality (R^2 or r)
Wearable microfluidic colorimetric sensor [11]	Vitamin C, H^+ (pH), Ca^{2+}, protein	Human tears	Analyte-induced color change in PDMS microfluidic patch captured as RGB signal for concentration mapping	Smartphone (RGB image capture)	CNN-GRU (1D for pH, 3D for others)	Image-to-concentration regression using CNN-GRU (Regression for Quantification)	Not reported	Not reported	0.001 (MAE)	$R^2 = 0.998$
Smartphone-based hydrogel colorimetric sensor [12]	Hyaluronidase (Haase) from bacteria	Clinical swabs, food	Hyaluronic acid (HA) degradation triggers CPRG release, reacts with β-galactidose and generates color changes	Smartphone (camera)	YOLOv5	Object detection and bacteria classification (Image Classification, Object Detection)	10 CFU/mL	92% (between gram + and gram -)	Not reported	$R^2 = 0.97$
Microfluidic plasmonic-enhanced colorimetric sensor [13]	H_2O_2	Cancer cell culture medium	Amplex Red reacts with H_2O_2 in presence of HRP, forming a pink dye; signal amplified by plasmonic nanostructures	Microscope (image capture)	Random Forest Classifier	Binary classification of H_2O_2 levels from RGB image (Image Classification)	1 picoMolar	91% (between high and low concentration classification)	Not reported	$R^2 = 0.98$
Lip-applied sensor [14]	pH	Skin surface (via lip application)	Anthocyaninl in lip pigment undergoes pH-triggered color shift captured via selfies	Smartphone (selfie camera)	CNN	Lip color classification into pH levels using CNN (Image Classification)	Not reported	92% (0.92	Not reported	Not reported
Multicolorimetric sensor array (AuNR-AgNP-based) (Plasmonic, Paper-based) [15]	HVA, VMA (tumor markers)	Human urine	Redox reaction between HVA/VMA and Ag^+ causes silver shell formation on Au nanorods, altering LSPR and generating multicolor shifts	Smartphone (RGB image)	PCA + LDA + PLSR	Multivariate regression and classification of tumor markers (Regression, Dim. Reduction + Classification)	0.22 µM (HVA) and 0.29 µM (VMA)	Not reported	100%	$R^2 = 0.999$ (HVA) $R^2 = 0.999$ (VMA)

Multiplexed Colorimetric Patch (PETAL) [16]	Temperature, pH, TMA, uric acid, moisture	Wound exudate (rat models)	Colorimetric sensors using liquid crystals, organic dyes, enzymes, and metal ions	Smartphone (patch image)	CNN	Image-based classification of wound biomarkers (Image Classification)	Not reported	94–96%	(blank)	(blank)
PDA-based lateral flow immunoassay (LFIA) (Lateral Flow) [17]	COVID-19 neutralizing antibody	Clinical serum	PDA-NPs conjugated with RBD antigen bind to antibodies; reduced PDA binding causes lighter test line; image processed via T/(T+C) grayscale ratio	Smartphone (test strip image)	Vision Transformer (ViT) + ResNet50	Band detection and antibody quantification using ViT (Regression, Object Detection)	160 ng/mL	Not reported	Not reported	Not reported
Dual-dye colorimetric RT-LAMP assay (Lateral Flow) [18]	SARS-CoV-2 RNA	Nasopharyngeal swabs	Isothermal amplification causes pH drop, triggering color change in Xylenol Orange and Lavender Green dyes; image analyzed post-reaction	Smartphone or camera (reaction tube image)	DETR-based model (ResNet50 + Transformer)	Tube segmentation and COVID result classification (Object Detection, Image Classification)	100% (reduced to 83% when diluted)	(blank)	(blank)	$R^2 = 0.998$
Paper-based multiplexed colorimetric biosensor (Paper-based) [19]	Cardiac and lipid biomarkers	Human serum	Targets (e.g., cTnI, HDL, LDL) separated and detected via electrophoresis-induced color change on paper	Scanner or smartphone (paper strip image)	CatBoost + PLS-DA, t-SNE (ensemble)	Color feature extraction and disease classification (Dim. Reduction + Clustering)	CtnI (1.210x10^-5 ug/mL) HDL (435.815 ug/mL) LDL (383.127 ug/mL)	75.2% for classification of acute myocardial infarction	Not reported	0.999, 0.9991, 0.999 respectively
Urinary disease colorimetric test array (Paper-based) [20]	Urinary disease markers	Human urine	Colorimetric reaction of multiple sensors (metal–organic complexes and chromogenic reagents) with urine	Smartphone (sensor array image)	Random Forest, SVM, kNN	Pattern classification of urinary markers (Image Classification)	Not reported	97% classification for UTI	(blank)	(blank)

			constituents captured via smartphone							
Colorimetric sensor using AuNPs[21]	Glucose	Urine samples	Glucose induces color change to AuNP	Smartphone	Image processing and illumination correction for accurate color interpretation across varying lighting conditions		Not reported	87.6% accurate glucose concentration prediction	Not reported	Not reported
Microfluidic sensor for artificial tears (Microfluidic) [22]	Glucose, cholesterol, pH	Synthetic tears	Gox/ChOx-mediated oxidation produces H_2O_2, catalyzing TMB color change via HRP; universal pH indicator used; smartphone captures RGB data	Smartphone (app-integrated µPAD images)	Deep Neural Network (DNN)	Regression for pH/glucose/cholesterol from artificial tear images (Regression for Quantification)	Glucose = 131 uM Cholestrol = 217 uM	100%	RMS=0.386	0.996 (glucose) 0.997 (cholesterol)
Sweat-based biosensor (Wearable) [23]	Glucose, pH, lactate	Human sweat	Chromogenic reactions triggered by sweat analytes across spatially arranged compartments; color changes recorded via smartphone	Smartphone (microfluidic chip images)	VGG16-based CNN	Color regression of sweat biomarker levels (Regression for Quantification)	Not reported	100% classification accuracy for all biomarkers in terms of quantity	(blank)	$R^2 = 0.9999$ for three biomarkers
HeLa cell-based metabolic colorimetric sensor [24]	Live HeLa cell viability (metabolic activity)	HeLa cell culture	pH-sensitive achromatic dye transitions (black to orange) based on cell density; saturation analyzed via smartphone images	Smartphone (achromatic saturation images)	Mask-RCNN	Quantification of live cell images (Image Classification)	51×10^4 cells	98%	(blank)	$R^2 = 0.959$

Colorimetric biogenic amine sensor for meat [25]	Biogenic amines	Chicken meat samples	Metal–azodye complex forming a fingerprint-based colorimetric response; analyzed via UV-vis absorbance and RGB imaging	UV–Vis scanner and smartphone (portable strip)	PCA, LDA, PLSR	Colorimetric amine pattern classification (Image Classification)	0.378 ppm (spermine)	Not reported	100% accuracy (cross validation), 83% for interference testing	Not reported
Tea polyphenol sensor during fermentation[26]	Tea polyphenols	Fermented green tea (w/ ultrasound)	RGB image extraction of CSA and multivariate calibration	Smartphone (RGB image of sensor array)	SVM	Regression and quality tracking for fermentation (Regression for Quantification)	Not reported	Not reported	Rc = 0.886, RMSEC = 0.042 mg/g, Rp = 0.862, and RMSEP = 0.043 mg/g	Not reported
Multiplexed troponin sensor (Nanozyme-based) [27]	Cardiac troponin I (cTnI)	Human serum	Cascade nanozyme-based colorimetric and photothermal signals from h-Prussian Blue in TMB-H_2O_2 system	Smartphone + thermometer (absorbance + thermal)	Artificial Neural Network (ANN, 3 hidden layers, 64 neurons)	Feature fusion from color and temperature signals for cTnI (Multimodal Fusion)	10.8 pg/mL	(blank)	(blank)	R^2 =0.9965
Sweat ion and pH patch sensor [28]	Na^+, K^+, pH	Human sweat during exercise	Printed chromogenic reagent zones and reference dye; color change recorded for in-situ analyte detection	Smartphone (sweat patch image with reference dye)	Explainable CNN (with ratiometric self-calibration)	Signal mapping for electrolyte and pH balance (Regression for Quantification)	classified and quantified with 100% accuracy	100% (≥50 nM)	(blank)	(blank)
Thiol-level cancer detection sensor [29]	Thiols (Cys, GSH, Hcy, DTT, MCH, TGA)	Standard solutions	Thiol-induced inhibition of metal ion–TPA@GQD nanozyme peroxidase-like catalysis of TMB-H_2O_2 reaction, creating distinct color patterns	UV–Vis reader or smartphone (RGB absorbance pattern)	Linear Discriminant Analysis (LDA)	Clustering of thiol-level profiles for disease classification (Dim. Reduction + Clustering)	50 nM thiol (not specified)	100% accuracy to separate and discriminate from different thiols	Not reported	Not reported

			(“fingerprints”) for LDA discrimination						
Biogenic amine sensor array [30]	Biogenic amines (tryptamine and spermine)	Meat and cottage cheese	Metal–azophenol complexes (C1–C11) respond to amines with colorimetric “fingerprint” patterns across 10 UV–Vis channels	Smartphone or UV–Vis scanner (sensor array image)	PCA, LDA	Color pattern recognition of food spoilage markers (Image Classification)	Tryptamine 0.40 ppm Histidine 0.42 ppm Spermine 0.45 ppm Spermidine 0.66 ppm	100% (LDA)	Not reported	R^2 = 0.96 (Tryp), 0.97 (Spermine)
Bimodal Visual Sensors Based on Mechanoluminescence and Biosensing [31]	Cariogenic bacteria (through pH from lactic acid)	Oral swabs, in vitro culture	Bacterial acid production (colorimetric pH shift via anthocyanin) and tooth pressure via mechanoluminescence	Smartphone (dual-mode mechanoluminescent + color image)	CNN-based model	Segmentation and bacterial profile analysis (Object Detection, Image Classification)	<1 mg/mL (estimated)	97.7% accuracy in the precise decoupling of visual signals	Not reported	Not reported
Heavy metal colorimetric sensor [32]	Cr^{3+}, Fe^{3+}, Al^{3+}, Ni^{2+}, Cu^{2+}, Zn^{2+}	Water and serum samples	AchE inhibition by metal ions alters enzymatic reaction with chromogenic substrate, producing color shift patterns	UV–Vis spectrophotometer (absorbance scan of arrays)	PCA	Metal concentration regression using pixel intensity (Regression for Quantification)	0.81 µM, 0.75 µM, 1.06 µM Cu2+, Cr3+, Al3+,	98% accuracy in p	(blank)	0.95,0.96,0.99 respectively
Tea authentication array sensor [33]	Tea polyphenols, adulterants	Tea infusion samples	TMB-H_2O_2 chromogenic system catalyzed by Bpy-Cu and Asp-Cu nanozymes; inhibition by polyphenols alters signal	Smartphone or scanner (nanozyme array image)	LDA, Decision Tree (DT), HCA	Classification of authentic vs adulterated tea via color (Image Classification)	Not reported	Discrimination accuracy was 100%.	Not reported	Not reported

AFB1 detection in ground peanut samples [34]	Aflatoxin B1 (AFB1)	Peanut extract (ground sample)	Aflatoxin B1 (AFB1) in food (e.g., peanuts)	Smartphone (fluorescent + colorimetric microneedle patch image)	ANN	AFB1 concentration prediction from patch image (Regression for Quantification)	0.6845 ng mL/1	(blank)	(blank)	$R^2 = 0.9974$
CO_2 strip colorimetric sensor [35]	CO_2	Ambient air (gas sample)	Color change induced by CO_2-mediated pH shift, captured as RGB ΔE across a 6-receptor array	Robotic camera and RGB sensor (automated platform)	Multi-target Bayesian Optimization (BO) integrated with robotic plcaatform	CO_2 level regression from colorimetric signal (Regression for Quantification)	400 ppm	(blank)	RMSE = 0.27%	(blank)
Paper-based glucose sensor [36]	Glucose	Human plasma	Enzyme-catalyzed colorimetric reaction using glucose oxidase (Gox) and horseradish peroxidase (HRP), with TMB for low and KI for high glucose concentration detection	Smartphone (flash/no-flash image pair)	Ensemble Bagging Classifier (EBC), Linear Regression	Glucose intensity prediction using Lab image values (Regression for Quantification, Classification)	Not reported	95% (TMB color indicator), 91% (KI color indicator)	(blank)	$R^2 = 0.97$ (high conc), $R^2 = 0.95$ (low conc)
Urine neurotransmitter sensor [37]	dopamine (DA), epinephrine (EP), norepinephrine (NEP), and levodopa (LD)	Human urine	Aggregation-based LSPR shift from AuNP interactions at different pH conditions	Smartphone (LSPR color shift image under pH variation)	LDA, PLSR	Catecholamine level estimation using feature-based models (Regression for Quantification, Classification)	0.3, 0.5, 0.2, and 1.9 mM for DA, EP, NEP, and LD,	100% (LDA)	(blank)	$R^2 = 0.99$ for 4 analytes
Smart μPAD for pH and glucose [38]	pH, Glucose	Aqueous lab-prepared solutions	For pH: Pani-NP undergoes EB to ES state transition; For glucose: Gox generates H_2O_2, reducing Pani-NPs, causing color shift (blue→green)	Smartphone (dipstick color image under ambient light)	RFR (best), DTR, SVR	RGB analysis for pH/glucose detection in μPADs (Regression for Quantification)	None reported	(blank)	(blank)	$R^2 = 0.96$ (pH), 0.92 (glucose)

Milk β-lactoglobulin strip (Lateral Flow)[39]	β-Lactoglobulin	Milk	Glucose-fueled EBFC for electrochemical + HRP/ABTS colorimetric detection using smartphone-assisted image processing	Smartphone (colorimetric + voltage strip readout)	Decision Tree (DT), Random Forest (RF), k-NN, SVM	Grayscale intensity detection for β-Lactoglobulin (Regression for Quantification)	0.0081 ng/mL,	93%	Not reported	Not reported
Albumin detection strip (Lateral Flow)[40]	Albumin	Urine	Protein concentration triggers color change on dipstick; captured by smartphone under varied lighting	Smartphone (dipstick image under varied lighting)	KNN classifier (vs RF, SVM)	Intensity ratio computation for albumin strip (Regression for Quantification)	4 mg/L	96%	Not reported	Not reported
H_2O_2 sensor (Spectrophotometric) [41]	Hydrogen peroxide	Exhaled breath	RGB signal mapping via colorimetric dye response (Eosin blue, $KmnO_4$, Starch-Iodine)	Smartphone (RGB mapping of breath test strip)	ANN Regression	Colorimetric pixel-based regression of H_2O_2 (Regression for Quantification)	0.011 ppm	94% accuracy for quantification		0.941
Saliva uric acid µPAD (Microfluidic) [42]	Uric Acid	Saliva	Prussian blue generation reaction with salivary UA forming blue complex	Smartphone (µPAD salivary test image)	Decision Tree Regressor (ML); Multiple Polynomial Regressor	Color space regression for uric acid quantification (Regression for Quantification)	Not reported	(blank)	MAE=4.2 ppm	Not reported

Table S2. Reported LOD values in AI-based studies, colorimetry only, colorimetry plus smartphone, and typical values in food, human, and environment samples

Analyte	**LOD AI-based (µM)***	**LOD colorimetry only***	**LOD colorimetry plus smartphone (no AI)***	**Typical values***
H2O2 in cells	0.000001 µM	10.24 µM[43]	0.24 µM[44]	0.01 µM[45]
VMA in urine	0.22 µM	0.340 µM	0.260 µM[46]	28.7 µM
HVA in urine	0.29 µM	0.313 µM	0.397 µM[46]	41 µM
COVID-19 antibody in human serum	1.07×10^{-6} (160 ng/mL)	6.00×10^{-5} µM[47] (9 ng/uL)	$2.11x10^{-10}$ uM	None
Cardiac troponin (cTnI) in human serum	5.06×10^{-10} (1.210 x 10^-5 ug/mL)	5.44×10^{-16} µM (0.013 pg mL−1)	1.63×10^{-8} µM 48 $3.9X10^{-4}$ µg/ml	0.02 ng/L[49]
LDL in human serum	1.277×10^{-4} µM (383.127 ug/mL)	7.33×10^{-10} µM[50] (2.1999 µg/mL)	1.77×10^{-8} [51] (5.31 mg/dl)	0.53[52] (100 mg/dL)
HDL in human serum	1.09 µM (435.815 ug/mL)	2 mg/dL[53] 5.00×10−8 µM	2.03×10^{-7} µM (8.10 mg/dl)[51]	40 mg/dL
Glucose in tears	131 uM (23.61 mg/L)	0.32 µM [54] 0.05765 mg/L	13.49 uM[55]	0.2 mM[56] (360 mg/L)
Cholesterol in tears	217 uM (83.87 mg/L)	1.9 µM[57] 0.7356 mg/L	0.00085 M	1.9 µM
Spermine in chicken meat	1.87 µM 0.378 ppm	0.57 uM[58] 0.115 mg/L	0.4644 uM 0.094 ug/mL[59]	988.4 uM 200 ppm[60]
Cardiac troponin (cTnI) in human serum	10.8 picogram/mL 4.52×10−16 µM	5.44×10^{-16} µM (0.013 pg mL−1)	$.63\times10^{-8}$ uM[48] $3.9X10^{-4}$ µg/ml	0.02 ng/L[49]
Tryptamine in meat	2.50 µM 0.40 ppm	20 nM[61] 0.0032 mg/L	1.74 µg/L[62] 1.086×10^{-5} uM	5 mg/kg meat[63]
Histidine in meat	2.71 µM 0.42 ppm	0.1 µM[64]	8 µg/L[65]	9.0 µM[66]
Spermine in meat	2.22 µM 0.45 ppm	0.57 uM[58] 0.115 mg/L	0.4644 uM 0.094 ug/mL[59]	988.4 uM 200 ppm[60]

**Values* were converted using molar masses

Table S3. Reported R^2 values in AI-based studies, colorimetry only, and colorimetry plus smartphone

Analyte	**LOD AI-based (μM)**	**LOD colorimetry only**	**LOD colorimetry plus smartphone (no AI)**
H2O2 in cells	0.998	0.9972	0.997
VMA in urine	0.999	0.996	0.997
HVA in urine	0.999	0.995	0.998
Cardiac troponin (cTnI) in human serum	0.999	0.990	0.981
LDL in human serum	0.999	0.9946	0.7917
HDL in human serum	0.999	0.9918	0.8018
Glucose in tears	0.996	0.994	0.995
Cholesterol in tears	0.997	0.993	0.993
Spermine in chicken meat	0.959	0.977	0.99209
Cardiac troponin (cTnI) in human serum	0.9999	0.990	0.981
Tryptamine in meat	0.9596	0.9969	0.987
Histidine in meat	0.952	0.9852	0.982
Spermine in meat	0.967	0.977	0.99209

*values correspond to cited studies in Table S2

Table S4. Statistical testing (Mann-Whitney U test) for LOD

Pair	**p-value**
AI-based vs Colorimetry + smartphone	p = 0.1610359
AI-based vs Colorimetry only	p = 0.3011529
AI-based vs Typical/Optimal	p = 0.2747575

*Statistical testing was conducted using R Studio's Wilcoxon rank-sum tests

Table S5. Statistical testing (Mann-Whitney U test) for R^2

Pair	**p-value**
AI-based vs Colorimetry + smartphone	p = 0.07539264
AI-based vs Colorimetry only	p = 0.1422894

*Statistical testing was conducted using R Studio's Wilcoxon rank-sum tests

Table S6. Data for bubble plot analysis with information obtained from Table S5

AI Subgroup	Analyte	Classification Accuracy	Analyte Group	Sensor Type
Object Detection	Hyaluronidase (HAase)	92	Enzymes / Enzyme Activity	Tube- or Well-Based Sensors
Traditional Machine Learning	Hâ,,Oâ,,	91	Small Molecule Metabolites	Tube- or Well-Based Sensors
Deep Learning	pH	92	Electrolytes / Ions	Wearable Patch Sensors
Ensemble / Hybrid Models	HVA	100	Tumor / Disease Biomarkers	Paper/Î¼PAD Sensors
Ensemble / Hybrid Models	VMA	100	Tumor / Disease Biomarkers	Paper/Î¼PAD Sensors
Transformer-based Models	SARS-CoV-2 RNA	100	Pathogen/Bacteria Detection	Tube- or Well-Based Sensors
Unspecified or Black-box	Cardiac and lipid biomarkers	75.2	Tumor / Disease Biomarkers	Paper/Î¼PAD Sensors
Traditional Machine Learning	Urinary disease markers	97	Tumor / Disease Biomarkers	Paper/Î¼PAD Sensors
Unspecified or Black-box	Glucose	87.6	Small Molecule Metabolites	Tube- or Well-Based Sensors
Deep Learning	Glucose	100	Small Molecule Metabolites	Tube- or Well-Based Sensors
Deep Learning	cholesterol	100	Small Molecule Metabolites	Tube- or Well-Based Sensors
Deep Learning	pH	100	Electrolytes / Ions	Tube- or Well-Based Sensors
Deep Learning	Glucose	100	Small Molecule Metabolites	Wearable Patch Sensors
Deep Learning	pH	100	Electrolytes / Ions	Wearable Patch Sensors
Deep Learning	lactate	100	Small Molecule Metabolites	Wearable Patch Sensors
Object Detection	Live HeLa cell viability	98	Tumor / Disease Biomarkers	Tube- or Well-Based Sensors
Dim. Reduction	Biogenic amines	100	Biogenic Amines / Spoilage Markers	Paper/Î¼PAD Sensors
Deep Learning	Naâ• º	100	Electrolytes / Ions	Wearable Patch Sensors
Deep Learning	Kâ• º	100	Electrolytes / Ions	Wearable Patch Sensors
Deep Learning	pH	100	Electrolytes / Ions	Wearable Patch Sensors
Dim. Reduction	Thiols	100	Composite / Multiplexed Panels	Tube- or Well-Based Sensors
Dim. Reduction	Biogenic amines	100	Biogenic Amines / Spoilage Markers	Tube- or Well-Based Sensors
Object Detection	Cariogenic bacteria	97.7	Pathogen/Bacteria Detection	Paper/Î¼PAD Sensors
Dim. Reduction	Heavy metals	98	Heavy Metals / Inorganics	Tube- or Well-Based Sensors
Dim. Reduction	Tea polyphenols	100	Polyphenols / Adulterants	Tube- or Well-Based Sensors

Ensemble / Hybrid Models	Glucose	95	Small Molecule Metabolites	Paper/Î¼PAD Sensors
Dim. Reduction	Catecholamines	100	Hormones / Neurotransmitters	Tube- or Well-Based Sensors
Traditional Machine Learning	Î²-Lactoglobulin	93	Macromolecules / Proteins	Tube- or Well-Based Sensors
Traditional Machine Learning	Albumin	96	Macromolecules / Proteins	Tube- or Well-Based Sensors
Deep Learning	Hâ,,Oâ,,	94	Small Molecule Metabolites	Tube- or Well-Based Sensors