

Article

Not peer-reviewed version

---

# Development and Validation of a Patient-Friendly Digital Assessment Platform for Precision Screening of Oral Anti-Obesity Medications (AOMs)

---

[Yawen Wang](#)<sup>\*</sup>, Jiaqi Chen, Rebekah Arias, Yingli Wang, Xiaoqing Yin

Posted Date: 16 January 2026

doi: 10.20944/preprints202601.1297.v1

Keywords: AOMs; precision medication; digital health; patient-friendliness; risk assessment



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Development and Validation of a Patient-Friendly Digital Assessment Platform for Precision Screening of Oral Anti-Obesity Medications (AOMs)

Yawen Wang <sup>1,\*</sup>, Jiaqi Chen <sup>1</sup>, Rebekah Arias <sup>1</sup>, Yingli Wang <sup>1</sup> and Xiaoqing Yin <sup>2</sup>

<sup>1</sup> Carnegie Mellon University, Pittsburgh, 15213, United States

<sup>2</sup> Stanford University, Stanford, 94305, United States

\* Correspondence: yawenw2@andrew.cmu.edu

## Abstract

This study developed a patient-facing digital health app for precision screening of oral anti-obesity medications (AOMs), integrating questionnaires, lifestyle factors, wearable signals, and prior metabolic indicators to enable personalized drug recommendations. Based on real-world testing data from 520 users, a weighted ensemble model was employed for multidimensional risk assessment, achieving 82% consistency between recommended outcomes and clinical prescriptions. UX testing validated the system's patient-friendliness, significantly enhancing appointment readiness and long-term adherence. This study provides a scalable tool for AOM clinical screening and follow-up.

**Keywords:** AOMs; precision medication; digital health; patient-friendliness; risk assessment

## 1. Introduction

Current clinical applications of AOMs lack precision screening mechanisms tailored to individual metabolic variations and behavioral characteristics. Significant heterogeneity in patient drug responses impacts intervention adherence and risk control efficiency. Integrating mobile sensing capabilities with structured risk assessment methods, developing a digital screening platform for patient self-assessment and remote assistance has become a necessary trend. To enhance recommendation accuracy and interface interaction performance, this study integrates multi-source health data to design a multi-modal integrated model deployable on edge devices. Systematic validation is conducted across system architecture, algorithm design, and user evaluation mechanisms.

## 2. System Architecture Design

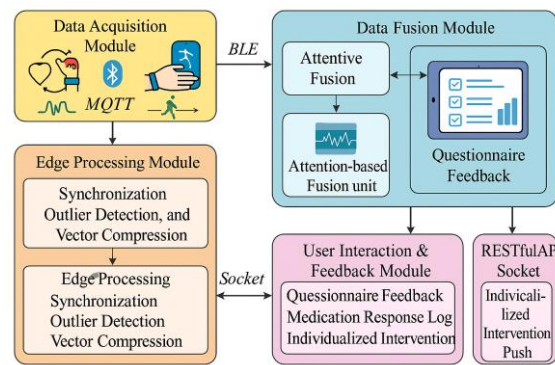
### 2.1. Overall Platform Architecture

Addressing the needs for adaptive precision screening and digital intervention evaluation in AOMs treatment, this platform adopts a four-layer system architecture (Figure 1), comprising data acquisition, edge processing, recommendation modeling, and user interaction layers. The data acquisition layer supports collecting physiological signals (optical glucose, EDA, HRV, gait, sleep) from wearable devices via BLE and MQTT protocols, enabling local caching and packet loss compensation. The edge processing layer hosts lightweight preprocessing modules on the Jetson Xavier NX to execute three core tasks: data synchronization, anomaly detection, and vector compression. The anomaly detection module applies a dual-rule strategy based on z-score ( $>3$ ) and interquartile range ( $IQR \times 1.5$ ). Readings that breach both thresholds within a 3-second window are flagged as anomalies. The system resolves these anomalies through three corresponding operations:

median replacement for isolated spikes (fewer than two consecutive points), cubic spline interpolation for gaps no longer than 1.5 seconds, and full exclusion of windows when anomalies exceed 20% of their length. This design ensures signal integrity, mitigates device-induced noise, and sustains an average processing latency of 41.3 ms[1]. Building on this preprocessed signal integrity, the platform advances into its modeling architecture, where a multi-stream fusion framework drives the AOMs recommendation task. The model leverages a residual gating strategy on the processing unit to enhance inter-modality attention learning. Its principal mapping logic is defined as:

$$\hat{y} = \sum_{i=1}^M \alpha_i \times f_i(x_i) \quad \text{Among them: } \alpha_i = \frac{\exp(\theta_i)}{\sum_{j=1}^M \exp(\theta_j)} \quad (1)$$

where  $x_i$  represents the feature of the  $i$  th modality,  $\alpha_i$  denotes its attention weight,  $\theta_i$  is the adaptive parameter, and  $f_i$  is the modality feature extraction function. The user interaction layer employs a front-end Flutter framework to design a responsive UX, supporting questionnaire feedback, medication reaction log uploads, and personalized intervention push notifications, ensuring the platform operates in a fully closed-loop manner.



**Figure 1.** Overall System Architecture Diagram.

## 2.2. Multi-source Data Integration Mechanism

To support refined modeling of individual drug responses within the AOMs screening platform, this system designs a multi-source heterogeneous data integration process covering four core data channels: questionnaire data, lifestyle behavioral characteristics, wearable physiological signals, and historical metabolic indicators. Data integration is implemented via a hybrid communication protocol stack, utilizing BLE for dynamic signal transmission and HTTPS channels for structured health record uploads [2]. All input data must meet unified format standards before entering the platform. The standardization process is modeled as follows:

$$\tilde{x}_{i,j} = \frac{x_{i,j} - \mu_j}{\sqrt{\sigma_j^2 + \varepsilon}}, \quad i \in [1, N], j \in [1, D] \quad (2)$$

Where  $\tilde{x}_{i,j}$  represents the  $j$  th feature dimension of the  $i$  th standardized sample,  $\mu_j, \sigma_j^2$  denotes the mean and variance of this dimension across training samples, and  $\varepsilon$  is the numerical stability constant. The platform deploys a unified data abstraction layer that interfaces with various data sources according to the input interface specifications in Table 1. This ensures dimensional consistency, sampling alignment, and noise suppression, while implementing data frame caching and interrupted transmission resumption through a streaming buffering mechanism [3].

**Table 1.** Interface Specifications and Protocol Support for Various Data Channel Connections.

Data Channel Type	Sample Dimension $D$	Access Protocol	Supported Devices/Sources	Data Frequency	Buffering Mechanism Description
Wearable Physiological Signals	15-dimensional (HR, EDA, temp)	BLE/MQTT	Garmin Vivosmart, Mi Band	5 Hz	5-second cycle buffering + packet loss compensation
Behavioral trajectory data	6-dimensional (step frequency, sedentary time, sleep/wake cycles, etc.)	BLE	Accelerometer + IMU	10 Hz	Linear interpolation within 10-second window
User questionnaire data	18 items (diet, exercise, medication, etc.)	HTTPS	Platform internal form module	Discrete collection	Fixed data structure, trigger-based collection and upload
Historical metabolic indicators	8 items (BMI, FBG, LDL, etc.)	HTTPS	ealth Information System (HIS) Interface	Offline import	Hospital API, CSV structure conversion integration

### 2.3. User Interaction and Interface Usability Strategy

To achieve structured collection of multi-source health information and proactive feedback management of patient drug reactions, a user interaction–interface adaptation system is established. This integrates multi-terminal UI design, layered response control mechanisms, and personalized interaction flows. The interaction architecture employs a Flutter+WebAssembly combination to build the frontend view layer, supporting unified rendering across iOS/Android/Web platforms. Dynamic loading adapts to terminal resolution and touch precision [4]. The backend utilizes asynchronous Socket channels and RESTful APIs to map user events (e.g., questionnaire submissions, drug reaction registrations) to task scheduling units, which then feed back to the recommendation engine to update preference vectors. Modular behavioral guidance mechanisms are embedded within interaction paths, configuring UI workflow blocks for different usage stages—including questionnaire completion, daily life logging, medication management, and individual feedback. All interaction modules dynamically load based on logical trigger conditions, employing component lazy activation and memory reclamation strategies to maximize response efficiency and load stability [5]. User-side operations are facilitated through multi-channel interface containers, forming a stable human-machine closed-loop input system.

## 3. Personalized Recommendation Algorithm Model Design

### 3.1. Feature Engineering and Data Preprocessing

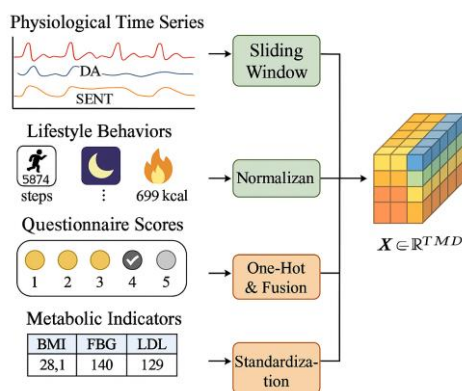
To support the high-dimensional risk modeling tasks of the multimodal AOMs personalized recommendation model, the system employs a unified tensor-based feature framework for reconstructing heterogeneous data structures, as illustrated in Figure 2. Input data comprises four major categories: physiological time-series signals (HR, EDA, skin temperature), lifestyle behavioral metrics (steps, sleep duration, energy expenditure), structured questionnaire scores, and historical

metabolic test values. For continuous modal data, a fixed-length sliding window strategy with a window size of 128 time steps and a stride of 32 is used to segment raw signals into third-order tensors[6].

To ensure methodological reproducibility, the normalization pipeline adopts explicit parameter settings. For each feature dimension, the mean and variance are pre-computed from the training cohort and stored as global statistics. A numerical stability constant of  $\varepsilon = 1 \times 10^{-5}$  is applied. Outlier suppression follows a bounded clipping strategy based on physiological priors: HR values below 40 bpm or above 180 bpm, EDA values outside  $[0, 15 \mu\text{S}]$ , and temperature outside  $[30^\circ\text{C}, 40^\circ\text{C}]$  are replaced using a 5-point linear interpolation window before normalization. The resulting normalized tensor representation follows Equation:

$$\tilde{x}_{t,m,d} = \frac{x_{t,m,d} - \mu_{m,d}}{\sqrt{\sigma_{m,d}^2 + \varepsilon}} \quad (3)$$

Among these,  $\tilde{x}_{t,m,d}$  represents the normalized tensor value across time step  $t$ , modal channel  $m$ , and feature dimension  $d$ .  $\mu_{m,d}$ ,  $\sigma_{m,d}^2$  denote the mean and variance of this dimension, respectively, while  $\varepsilon$  is the numerical stability constant. Discrete questionnaire features adopt univariate encoding combined with a weighted Likert aggregation rule using weights  $w = \{0.8, 1.0, 1.2\}$  for lifestyle, medication history, and symptom items respectively. Historical metabolic indicators are normalized using z-score scaling with outlier trimming at  $\pm 3\sigma$ , ensuring stable gradient behavior during model training.



**Figure 2.** Tensor Encoding Structure for Multimodal Input Data.

### 3.2. Multimodal Data Fusion and Weighted Ensemble Model

To enable multidimensional risk inference tailored for AOMs interventions, we designed a weighted ensemble recommendation model architecture. Built upon multimodal input tensors  $\mathbf{X} \in \mathbb{R}^{N \times M \times D}$ , it achieves structural fusion of heterogeneous data through modality-aware attention mechanisms. The model comprises three sub-networks: time-series branch, static structure branch, and logical scoring branch [7]. Each branch independently extracts modal semantics before undergoing risk prediction through a weighted ensemble layer. The fusion formula is as follows:

$$\hat{y} = \sum_{k=1}^K w_k \times f_k(\mathbf{X}_k), \text{ with } w_k = \frac{\exp(\alpha_k)}{\sum_{j=1}^K \exp(\alpha_j)} \quad (4)$$

where  $\hat{y}$  represents the final recommendation risk score,  $\mathbf{X}_k$  denotes the input from the  $k$ th modal sub-input,  $f_k(\cdot)$  is the output function of the  $k$ th branch,  $\alpha_k$  is the modal attention score, and  $w_k$  is the corresponding normalized weight. The integration layer employs a residual fusion structure with a Sigmoid gating mechanism to control redundant information transformation between modalities.

### 3.3. Model Training and Validation Methods

To ensure the AOMs recommendation model exhibits stable convergence and generalization capabilities under multimodal input scenarios, an end-to-end deep fusion framework is adopted, combined with supervised risk labels for target fitting. The training dataset comprises structured data from 520 patients, with input tensor dimensions of  $\mathbf{X} \in \mathfrak{R}^{N \times M \times D}$  and binary classification labels (recommended/not recommended). The model training objective is defined as a weighted combination of cross-entropy loss and L2 regularization:

$$L_{total} = -\sum_{i=1}^N y_i \log(\hat{y}_i) + \lambda \sum_{k=1}^K \|\theta_k\|^2 \quad (5)$$

where  $y_i$  represents the true label,  $\hat{y}_i$  denotes the model's predicted probability,  $\theta_k$  is the learnable parameter of the  $k$  module, and  $\lambda$  is the regularization strength coefficient. The Adam optimizer was employed during training with an initial learning rate of 1e-3, a batch size of 32, and 100 training epochs. A five-fold cross-validation structure was adopted for validation, automatically partitioning each training round into an 80% training set and a 20% validation set for supervised updates [8].

## 4. Platform Validation and Performance Evaluation

### 4.1. Recommendation Consistency Evaluation

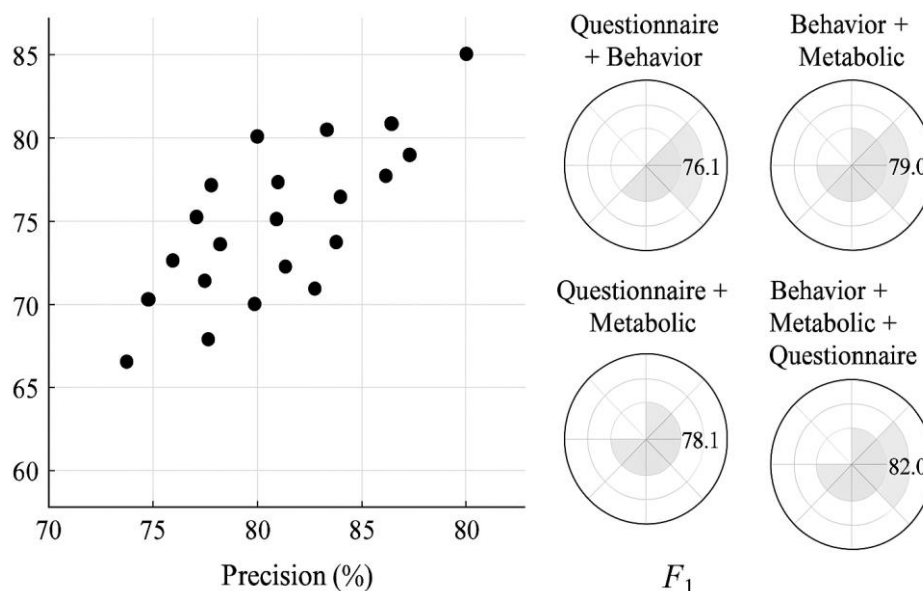
To validate the consistency between the AOMs recommendations generated by the platform based on multi-source health data fusion and actual clinical prescriptions, this study conducted comparative experiments using data from 520 users with complete electronic medical records and follow-up labels. During the experiment, the system input consisted of standardized multimodal tensor data. Labels were calibrated based on the joint prescription opinions of two endocrinology attending physicians at the medical institution, with recommendation results defined as a binary output (recommended/not recommended) [9]. Consistency evaluation utilized Top-1 accuracy (Acc), match rate (Match), and F1 score metrics, defined as follows:

$$\text{Match} = \frac{1}{N} \sum_{i=1}^N I[\hat{y}_i = y_i], \quad F1 = \frac{2 \times P \times R}{P + R} \quad (6)$$

where  $\hat{y}_i$  represents the system's output recommendation label,  $y_i$  denotes the actual clinical label,  $I[\cdot]$  is the indicator function,  $P$  is precision, and  $R$  is recall. Results are shown in Figure 3. The "Behavior + Metabolic + Questionnaire" combination achieved the best performance among all modality configurations, with Precision, Recall, and F1 reaching 83.3%, 80.8%, and 82.0% respectively—significantly higher than other combinations. This indicates stronger generalization capabilities in integrated information perception and risk discrimination. To further assess the relative effectiveness of the weighted ensemble model, we introduced two mainstream classification models—Random Forest (RF) and XGBoost—as performance baselines under the same training dataset and modal tensor input. These models were trained using default hyperparameters and five-fold cross-validation to ensure fair benchmarking. As shown in Table 2, while RF and XGBoost achieved acceptable accuracy (F1 scores of 76.4% and 78.1% respectively), the weighted ensemble model outperformed both in precision (83.3% vs. 74.5%/75.8%) and recall (80.8% vs. 76.2%/78.5%), indicating its superior capability in capturing latent feature interactions across modalities. These results validate the enhanced adaptation performance and clinical decision-making reliability of our fusion model.

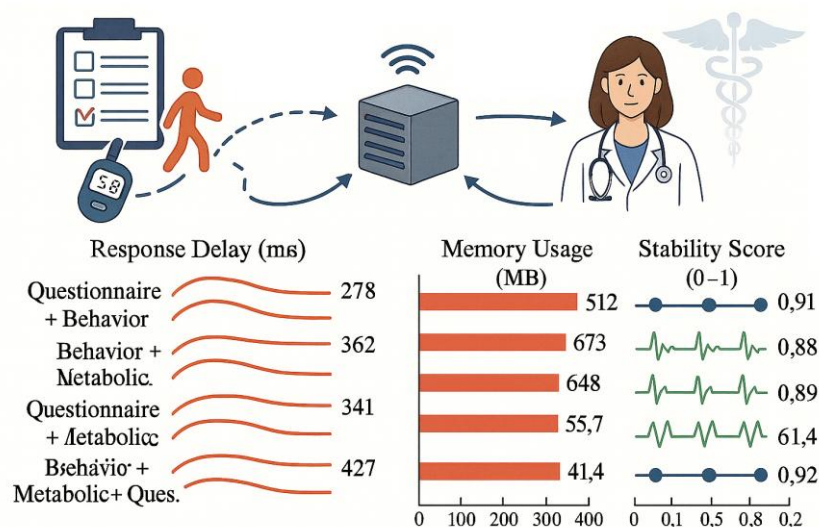
**Table 2.** Comparison of Recommendation Performance between Models.

Model Type	Precision (%)	Recall (%)	F1 Score (%)
Weighted Ensemble	83.3	80.8	82.0
Random Forest (RF)	74.5	76.2	75.3
XGBoost	75.8	78.5	77.1

**Figure 3.** Comparison of Recommendation Consistency Metrics.

#### 4.2. System Response and Resource Overhead Analysis

To evaluate the system's practical deployment performance in mobile scenarios, the platform selected Jetson Xavier NX (6-core CPU + 384-core CUDA) as the edge processing device, simulating real-world request scenarios based on user data. Tests used four modality combinations as input conditions, uniformly invoking the platform's model service interface. The complete workflow latency—from data upload, preprocessing, model inference to recommendation return—was progressively recorded while monitoring GPU utilization and RAM dynamic allocation [10]. Under all modality combinations, the platform's overall response latency remained below 600ms. The "Questionnaire+Behavior" combination achieved the fastest response time of 278ms due to simplified preprocessing logic. Despite its complex computational path, the three-modality fusion configuration maintained a response time of 427ms through structural reuse and cache optimization, with memory consumption at 721MB—significantly below the platform's upper threshold (1.5GB). Figure 4 presents key performance metrics across different modality combinations, including inference duration, average resource utilization, and system stability scores. Results demonstrate the platform's strong mobile adaptation capabilities and computational scheduling efficiency, making it suitable for edge deployment in AOM screening applications.



**Figure 4.** Comparison of System Resource Overhead and Response Performance Across Multimodal Combinations.

#### 4.3. UX Testing and Patient Usability Evaluation

To quantify the platform's interaction friendliness and usability in real-world scenarios, this study designed multi-dimensional UX testing tasks based on the Nielsen Usability Evaluation Framework. These tasks covered eight core interaction points, including system login, questionnaire completion, recommendation viewing, and feedback submission. The test subjects comprised 52 obese users with genuine medication needs. All participants were pre-assessed using a simplified digital literacy index (DLI) and categorized into three groups (low, medium, high) based on operational experience with smartphones and healthcare apps. Among them, 65.4% belonged to the medium DLI group, and 71.2% were aged between 30 and 45 years, indicating a relatively tech-familiar demographic. Process data were collected using a tripartite approach of "task guidance + time logging + operation observation." Subjective experiences were quantified using the NASA TLX scale and the SUS system usability rating system. The platform achieved an overall UX score of 86.7/100, with a task completion rate exceeding 96% and an average response time per task maintained below 3.8 seconds. Subgroup analysis revealed marginally higher usability scores in the high-DLI group (avg. 89.1/100) compared to the low-DLI group (avg. 82.3/100), suggesting a moderate influence of user familiarity on interaction fluency. While the "Individualized Recommendation View" module achieved the highest satisfaction score (92.4/100), further testing in digitally underrepresented populations (e.g., elderly users and low-DLI groups) will be necessary to validate generalizability. Future iterations will incorporate stratified sampling and inclusive UI adjustment strategies to enhance robustness across user segments.

## 5. Conclusions

The platform constructs a risk closed-loop model by integrating wearable signals, behavioral data, and questionnaire structural features. Through fusion optimization and multimodal mechanisms, it effectively enhances recommendation consistency and interactive responsiveness. Validation results demonstrate the system's high model accuracy, operational efficiency, and user-friendliness, making it suitable for screening scenarios tailored to individual patient characteristics. Future applications may extend to other metabolic intervention decision domains. Incorporating cross-platform modeling, dynamic feedback mechanisms, and model fine-tuning strategies could enhance generalization capabilities and long-term clinical applicability.

## References

1. Zheng, X., Dwyer, V. M., Barrett, L. A., Derakhshani, M., & Hu, S. (2023). Rapid vital sign extraction for real-time opto-physiological monitoring at varying physical activity intensity levels. *IEEE Journal of Biomedical and Health Informatics*, 27(7), 3107-3118.
2. Beniwal S S, Lamo P, Kaushik A, et al. Current status and emerging trends in colorectal cancer screening and diagnostics[J]. *Biosensors*, 2023, 13(10): 926.
3. Fahrner L J, Chen E, Topol E, et al. The generative era of medical AI[J]. *Cell*, 2025, 188(14): 3648-3660.
4. Makram T S, Eid S M, Abu-Dahab M, et al. Formulation of Saxagliptin Oral Films: Optimization, Physicochemical Characterization, In-Vivo Assessment, and In-Vitro Real-Time Release Monitoring via a Novel Polyaniline Nanoparticles-Based Solid-Contact Screen Printed Ion-Selective Electrode[J]. *AAPS PharmSciTech*, 2024, 25(5): 116.
5. Ghosh A, Sanyal A, Kumari M, et al. Innovations in additive manufacturing approaches for the fabrication of advanced drug delivery platforms, biomimetics, and predictive 3D microphysiological interfaces[J]. *Journal of Materials Chemistry B*, 2025, 13(38): 11928-11970.
6. Behera P P, Kumar N, Kumari M, et al. Integrated microfluidic devices for point-of-care detection of bio-analytes and disease[J]. *Sensors & diagnostics*, 2023, 2(6): 1437-1459.
7. Los A, Goldsmith G, Kilty R L H, et al. PCR52 Use of Patient Reported Outcomes in Key Trials Supporting Marketing Authorisation: 5-Year Analysis of Company Evidence Submitted for National Institute for Health and Care Excellence (NICE) Single Technology Appraisal (STA)[J]. *Value in Health*, 2023, 26(12): S458-S459.
8. Huang L, Luo S, Tong S, et al. The development of nanocarriers for natural products[J]. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology*, 2024, 16(3): e1967.
9. Elhabal S F, Mohammed Ali E, Hababeh S, et al. Harnessing Nature for Breast Cancer Management: Effects of Fisetin-Loaded Nigellasomes Embedded in Microneedles Improve Tumor Suppression and Reduce Oxidative Stress[J]. *Pharmaceutics*, 2025, 17(11): 1392.
10. Zu M, Ma Y, Zhang J, et al. An oral nanomedicine elicits in situ vaccination effect against colorectal cancer[J]. *ACS nano*, 2024, 18(4): 3651-3668.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.