Evidence-Based Study on Low-Burden Digital Phenotyping for Precision Screening of Oral Anti-Obesity Drug Efficacy

Yawen Wang * , Xiaoqing Yin , Jiaqi Chen , Yingli Wang

Posted Date: 23 March 2026

doi: 10.20944/preprints202603.1764.v1

Keywords: AOMs; low-burden digital phenotyping; patient-side app; interpretable modeling; time-series prediction; adherence

 doi:10.20944/preprints202603.1764.v1

Article

Evidence-Based Study on Low-Burden Digital Phenotyping for Precision Screening of Oral Anti-Obesity Drug Efficacy

Yawen Wang [1,*], Xiaoqing Yin [2], Jiaqi Chen [1] and Yingli Wang [1]

[1] Carnegie Mellon University, Pittsburgh, 15213, United States

[2] Stanford University, Stanford, 94305, United States

* Correspondence: yawenw2@andrew.cmu.edu

Abstract

To enable scalable and accurate early screening of oral anti-obesity medications (AOMs), we propose a low-burden digital phenotyping framework based on long-term patient-submitted data, including step count, heart rate, sleep summary, short diet frequency, and patient-reported outcomes (PROs). A cohort of 400 overweight or obese individuals was monitored for 24 weeks. A "behavioral rhythm–metabolic risk" feature set, comprising evening activity ratio, sleep consistency, postprandial subjective hunger, and mood/stress entries, was constructed. The hierarchical GBT + TFT model predicted 12-week weight loss ≥5% with high accuracy, achieving an AUC of 0.88 on external time-window validation and 0.83 using only the first 14 days of data, supporting early screening and rapid feedback. This study demonstrates that low-burden digital phenotyping can serve as a core capability for precision screening of AOMs and provides a practical pathway for patient-side applications and stratified follow-up.

CCS CONCEPTS:

Applied computing~Life and medical sciences~Health informatics

Keywords: AOMs; low-burden digital phenotyping; patient-side app; interpretable modeling; time-series prediction; adherence

1. Introduction

Oral anti-obesity medications (AOMs) demonstrate substantial variability in therapeutic response across individuals. Conventional screening and evaluation strategies primarily rely on randomized clinical trials or prolonged follow-up periods, where treatment response is assessed through longitudinal monitoring of weight loss and metabolic indicators. Although these approaches provide rigorous clinical evidence, they are resource-intensive and difficult to scale in real-world settings. As a result, they are not well suited for rapid, early-stage identification of individuals who are most likely to benefit from pharmacological interventions. Developing a data-driven method that can identify potential responders at an early stage of treatment has therefore become an important challenge in precision obesity management.

Recent advances in digital health technologies have introduced low-burden digital phenotyping as a promising approach for continuous health monitoring. Through wearable devices and mobile health platforms, behavioral and physiological signals—such as daily step count, sleep patterns, heart rate dynamics, dietary habits, and self-reported mood—can be collected in a longitudinal and unobtrusive manner. These data streams capture individuals' behavioral rhythms and metabolic risk profiles with high temporal resolution and strong ecological validity. Leveraging these characteristics, digital phenotyping offers a scalable pathway for early prediction of treatment outcomes. In this study, we construct a behavioral rhythm–metabolic risk feature set and develop a

 doi:10.20944/preprints202603.1764.v1

hierarchical predictive framework that integrates Gradient Boosting Trees (GBT) with a Temporal Fusion Transformer (TFT). By combining interpretable feature selection with advanced time-series modeling, the proposed framework aims to improve prediction accuracy while maintaining interpretability, thereby providing a scalable analytical tool to support early identification of patients who are most likely to respond to AOM therapy.

2. Building a Low-Burden Digital Phenotyping Data System

Given the limited long-term retention rate of patient-side applications in real-world settings, the data system was designed following the principle of "long-term self-collection with minimal interaction." This approach aims to reduce user burden while maintaining continuous behavioral monitoring. The system performs unified time alignment and quality control across both wearable-device data streams and self-reported inputs to ensure temporal consistency and comparability among heterogeneous data sources [1]. Step count and heart rate are collected at one-minute intervals and locally aggregated into five-minute summaries before transmission, which effectively reduces data transmission load and storage requirements while preserving meaningful temporal patterns.

Sleep data are restricted to nightly summary fields, including sleep onset time, wake time, duration of sleep stages, and overall sleep efficiency. Daytime recording for sleep modules is disabled to prevent redundant or irrelevant entries. Dietary information is captured through rapid meal-frequency check-ins lasting approximately 3–5 seconds, using fixed timestamps combined with post-meal hunger ratings on a 0–10 scale. This design allows the system to approximate short-term satiety dynamics and eating behavior without imposing heavy reporting demands. Patient-reported outcomes (PROs) include 4–6 Likert-scale items—such as mood or perceived stress—collected once daily to provide additional contextual information regarding psychological state and behavioral variability.

The preprocessing pipeline consists of several steps. (1) Missing value screening and imputation: variables with more than 30% missing data within the aggregation window are removed, while remaining missing values are imputed using a combined forward-fill and 7-day median backfill strategy. Observed missing rates were relatively low, ranging from 0–5% for step count, 0–3% for heart rate, 5–7% for sleep summaries, and 0–2% for PRO items. (2) Abnormal heart-rate correction, where values outside the physiological range of 40–180 bpm are clipped to reduce the impact of sensor artifacts. (3) Cross-device offset calibration using the formula $HR_{calibrated}=HR_{raw}+\Delta HR_{device}$, where ΔHR_{device} is derived from simultaneous measurements across different devices. (4) A seven-day rolling offset is applied to smooth diurnal rhythms and stabilize temporal patterns. Together, these procedures produce standardized time-series sequences suitable for computing the Behavioral Rhythm–Metabolic Risk features used in subsequent predictive modeling [2].

3. Time-Series Modeling Methods for Predicting AOM Efficacy

3.1. Modeling the Oral Anti-Obesity Medication Efficacy Prediction Problem

To support the timely availability of AOM efficacy evaluation in patient-oriented applications, time-series modeling transforms low-burden digital phenotypic sequences collected from baseline through T into probabilistic predictions of the 12-week response label. Treatment response is defined using a clinically feasible threshold:$y=I(\Delta W12/W0\leq -0.05)$, where $\Delta W12=W12-W0$. Here, W0denotes baseline body weight and W12 represents body weight at week 12. This definition corresponds to the widely accepted clinical benchmark of at least 5% body weight reduction, which serves as a reliable supervision signal for early efficacy prediction models.

① Input Design: A multivariate sequence of length $T\in\{7,14,21,28\}$is constructed using daily observations, including step count, resting heart rate, sleep efficiency, sleep onset time, night activity ratio, meal frequency, post-meal hunger score (0–10), and mood/stress scale. These variables

 doi:10.20944/preprints202603.1764.v1

collectively reflect behavioral, physiological, and psychological aspects of patient health, enabling the model to capture multidimensional lifestyle patterns that may influence medication response.

② Missing Data Handling: As wearable-device data and self-reported records are often incomplete, a hybrid imputation strategy combining Forward Fill and 7-Day Median Backfill is applied. In addition, a missing mask mtm_tmt is recorded and incorporated as an auxiliary modeling feature, allowing the model to explicitly learn the structure of missingness and reduce bias caused by data sparsity.

③ Learning Objectives: The model is trained using a calibrated binary classification loss, with class weights adaptively determined according to the proportion of positive responses to mitigate class imbalance. The predicted probabilities are further calibrated using temperature scaling to improve probabilistic reliability. These calibrated risk probabilities can then be directly used in downstream modules for risk stratification and clinical interpretation [3]. Through this framework, the model not only supports early prediction of treatment efficacy but also provides quantitative evidence for personalized intervention and monitoring strategies.

3.2. GBT-Based Static Behavior-Metabolic Risk Feature Modeling

In order to compress daily sequences of low-burden digital phenotypes into a robust learnable "static behavioral-metabolic risk" representation, we first carry out multiscale aggregation at the previous T-day window (T ∈ {7,14,21,28}) [4], producing a static vector xs: containing the mean/quantile values of the steps (p10, p50, p90), the median resting HR and MAD, the average sleep efficiency and consistency (SDT, SD), the percentage of nighttime activity (Steps 18: 00 – 24: 00/Total Daily Steps), Average Meal Frequency, Mean and Slope of Postmeal Hunger (0 – 10 Scale), Mean and Slope, Mean and Slope of Average and Average, Missing Data Hiding Ratio. The heavy-tailed variables were trimmed with log (1 + x) and winsorization (1%/99%) .GBT used a binary classifier with parameters: the number of trees N ∈ [300,800], the maximal depth d ∈ [3,8], the learning rate η ∈ [0.02,0.1], row sampling 0.6 – 0.9, minimum leaf node sample size 20 – 80. Stratified cross-validation with time windows was employed to prevent information leakage. For the TFT, the output risk prior was used as a static context input, whereas the SHAP contributions were kept for individual interpretation:

$$\hat{p} = \sigma(f_{\mathrm{GBT}}(x_s;\Theta)) \tag{1}$$

Where $\hat{p}$ represents the 12-week response probability estimate, $\sigma(\cdot)$ denotes the Sigmoid mapping, $f_{\mathrm{GBT}}(\cdot)$ is the boosting tree ensemble function determined by parameters Θ (tree structure and leaf weights), andxs` is the static feature vector5.obtained through aggregation within window T and quality control.

3.3. Multivariate Time-Series Behavioral Feature Prediction Model Based on TFT

Under the constraint of static GBT prior $\hat{p}$, TFT characterizes the dynamic evolution of treatment response risk from multi-variable daily sequences sustainably collected via patient-side apps.The model jointly encodes input sequences of length T∈{7,14,21,28} $X_{1:T} \in \mathrm{R}^{T\times F}$ with missingness masks $M_{1:T} \in \mathrm{R}^{T\times F}$, where F covers step count, resting heart rate, sleep efficiency and onset time, nighttime activity proportion, meal frequency, post-meal hunger (0–10), mood/stress entries, and their 7-day rolling statistics;The gate variable selection network sparsely weights static context $s = [\hat{p}, x_s]$ and time-varying features6 . Subsequently, LSTM local sequence modeling and multi-head attention capture cross-day compliance fluctuations and rhythm drift, outputting temperature-calibrated risk probabilities for rolling updates in the app:

$$\hat{y} = \sigma(W_o \mathrm{Attn}(\mathrm{LSTM}(X_{1:T}, M_{1:T}), s) + b_o) \tag{2}$$

 doi:10.20944/preprints202603.1764.v1

Where $\hat{y}$ represents response probability output, $\sigma(\cdot)$ denotes Sigmoid, W_o, b_o signifies output layer parameters, $\mathrm{LSTM}(\cdot)$ generates temporal hidden states, $\mathrm{Attn}(\cdot)$ is the multi-head attention fusion module, and s denotes the context vector concatenated from GBT risk prior and static features. To visually illustrate rolling prediction outputs, Figure 1 displays probability maps updated daily.

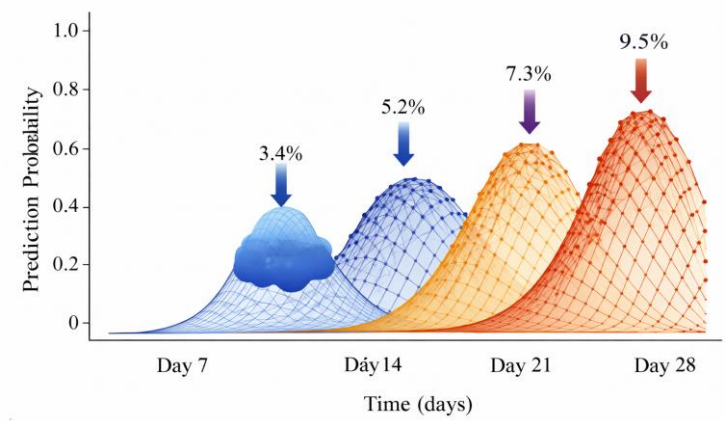

Figure 1. Output Pattern Diagram.

3.4. Construction of the GBT + TFT Fusion Hierarchical Prediction Framework

In order to steadily integrate low-burden digital phenotypes into the accurate screening procedure of oral anti-obesity drugs, the fusion framework uses a "static layer + dynamic correction" structure, as illustrated in Fig. 2. ① Static Layer: Aggregation of behavioral and metabolic risk vectors from T = 7/14/21/28. GBT generates risk priors and feature importance within ranges N = 300 – 800, d = 3 – 8, η = 0.02 – 0.1 to coarsely classify high/middle/low risk cohorts and provide interpretable anchors; ② Dynamic Layer: TFT receives daily granularity multivariate sequences and missing information masks. With attention heads: h = 4/8, hidden dimensions 64/128, and dropout 0.1 – 0.3, it performs temporal recalibration on coarse-grained results and outputs rolling probabilities; ③ Fusion Rules: Utilizes GBT priors as the TFT static context, performs temperature calibration and threshold locking on the output, providing daily updates on the application side and controllable computing on the device [7].

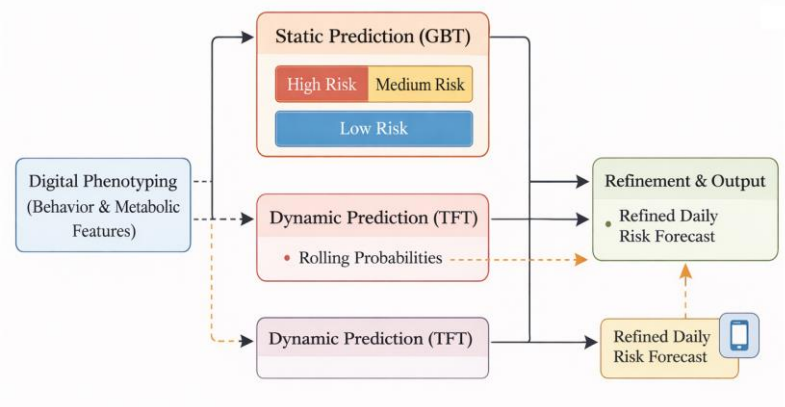

Figure 2. GBT+TFT Layered Fusion Framework Diagram.

Stratified time-window cross-validation with five folds was employed to prevent information leakage, ensuring each fold preserved the positive/negative response ratio while maintaining temporal order. GBT hyperparameters, including the number of trees N ∈ [300,800], maximum depth d ∈ [3,8], learning rate η ∈ [0.02,0.1], row sampling 0.6–0.9, and minimum leaf sample size 20–80, were optimized via grid search. TFT hyperparameters, including attention heads h ∈ {4,8}, hidden dimensions 64/128, and dropout rates 0.1–0.3, were similarly tuned using grid search on the training

 doi:10.20944/preprints202603.1764.v1

folds. The static GBT output was used as a context vector for TFT, and SHAP contributions were preserved for individual-level interpretation. The GBT + TFT fusion hierarchical framework thus integrates static risk priors with dynamic sequence modeling, providing daily rolling risk probabilities for patient-side applications while maintaining interpretable feature contributions.

3.5. Explainable Modeling Approach and Individual Contribution Analysis Mechanism

In order to ensure that the low burden digital phenotypes can be used jointly by clinicians and applications in the accurate screening of the efficacy of oral anti-obesity drugs, the explanation mechanism is implemented as follows: "Static Tree Model Interpretation + Temporal Attention Attribution + Individual Report Generation"8 . For the GBT layer, SHAP consistency explanations decompose individual static feature vectors x_s into baseline log-likelihood and feature contributions:

$$\log it(\hat{p}) + \phi_0 + \sum_{j=1}^{d} \phi_j \tag{3}$$

Where $\hat{p}$ represents the GBT output probability, $\log it(\cdot)$ denotes the log-odds transformation, ϕ_0 signifies the overall baseline term, d indicates the static feature dimension, and ϕ_j represents the marginal contribution of the j th feature (e.g., average sleep efficiency, proportion of nighttime activity, average post-meal hunger, emotional fluctuations) to the individual's prediction.For the TFT layer, we introduce dual time-variable attribution using attention weights $a_{t,k}$ to characterize the relative impact of the t th time-varying variable k on the output, normalized to ensure cross-window comparability:

$$\tilde{a}_{t,k} = \frac{\exp(a_{t,k})}{\sum_{t'=1}^{T}\sum_{k'=1}^{F}\exp(a_{t',k'})} \tag{4}$$

where T represents the input day window, F denotes the number of time-varying variables, and $\tilde{a}_{t,k}$ indicates the normalized importance. To generate auditable individual contribution summaries, static SHAP and temporal attribution are fused into a unified interpretability score c_k . Missingness masks $m_{t,k} \in \{0,1\}$ are penalized to prevent "missing-driven explanations":

$$c_k = \sum_{t=q}^{T} \tilde{a}_{t,k} x_{t,k} - \lambda \sum_{t=1}^{T} (1 - m_{t,k}) \tag{5}$$

where $x_{t,k}$ denotes the observed or imputed input value of variable k on day t , λ is the missingness penalty coefficient (set between 0.05–0.2 for endpoint stability), and c_k is used to output "Top-K Interventions + Confidence Markers", which are included in the individual interpretation records along with the drug adherence fields (dose control rate, missed dose interval days, continuous treatment duration, etc.), are used to support stratified monitoring as well as to tune the parameters of the threshold strategy [9].

4. Experimental Results and Analysis

4.1. Study Population and Experimental Dataset Construction

In order to empirically assess low-burden digital phenotypes in the precision screening of the efficacy of oral AOM, 400 overweight/obese individuals who completed 24 weeks of follow-up were included in the experimental data set, generating a combination of "daily time series + weekly weight" samples. The Behavioral Self Report Component gathered short diet frequency records and PROs

 doi:10.20944/preprints202603.1764.v1

(postprandial subjective hunger scores, mood/stress items, etc.), which were combined with wearable data in the same timetime.Weight was used as a monitored signal recorded once a week, with a response tag defined as 5% or more for 12 weeks. A preliminary screening subset is constructed using only the first 14 days of sequences for device-side usability testing. In order to verify the reliability of the time transmission, the training/validation sets are divided by an external time window, which ensures the assessment [10].

4.2. Performance Evaluation of Low-Burden Digital Phenotyping Model Predictions

We systematically evaluated the performance of the proposed low-burden digital phenotype framework in screening the effectiveness of oral anti-obesity drugs.In order to verify the early prediction ability of the model in real-world patient side application scenarios, different input time windows (7,14,21,28 days before) and external time windows were established. AUC, F1-SCORE, and recall were used as the primary measure of response for 12 weeks, with a weight loss of more than 5% at 12 weeks. The detailed results are shown in Table 1.

Table 1. Comparative Predictive Performance of Different Model Configurations on External Validation Set.

Model Configuration	Input Window	AUC	Precision	Recall	F1-score
GBT Static	14 days	0.81	0.76	0.72	0.74
TFT Timing	14 days	0.85	0.79	0.77	0.78
Fusion Model	7 days	0.79	0.73	0.7	0.71
	14 days	0.88	0.84	0.82	0.83
	21 days	0.89	0.85	0.83	0.84
	28 days	0.9	0.86	0.84	0.85

The proposed GBT + TFT hierarchical fusion framework, as illustrated in Table 1, achieves an AUC of 0.88 with a 14 day input window. Compared with the separate GBT (0.81) and the TFT time model (0.85), this is an improvement of 8.6% and 3.5%, respectively, which fully demonstrates the complementary benefits of the static behavior-metabolic risk priors and the temporal dynamic correction.Notably, the AUC of the model remains at a usable level of 0.79 when the input window is reduced to 7 days. However, when the window is prolonged to 21 and 28 days, the performance gains gradually plateau (the AUC increases 0.01 – 0.02). This provides quantitative evidence for the implementation of "early screening and quick feedback" in patient-side applications — confirming that the 14 day collection cycle is an optimal balance between performance and retention. Based on the accurate recall balance, the fusion model achieves a precision of 0.84 and a recall of 0.82 at 14 days, with an F1 score of 0.83. This demonstrates a strong ability to identify positive and negative cases, effectively supporting a clinical stratified screening needs.Furthermore, the efficacy of the missing data treatment strategy has been verified: over a 14 day period, the version that includes the missing value masking in the modeling results in an AUC 0.03 greater than that of the simpler mean-filling version. These results provide a basis for the study of interpretation and compliance interaction.

4.3. Analysis of the Impact of Different Time Window Data on Predictive Performance

In order to further explore the effect of time window length on early screening effectiveness, the AUC box plot for the ensemble model was plotted on the basis of a time window stratified 5 times cross validation (7, 14, 21, 28 days), as illustrated in Figure 3.

 doi:10.20944/preprints202603.1764.v1

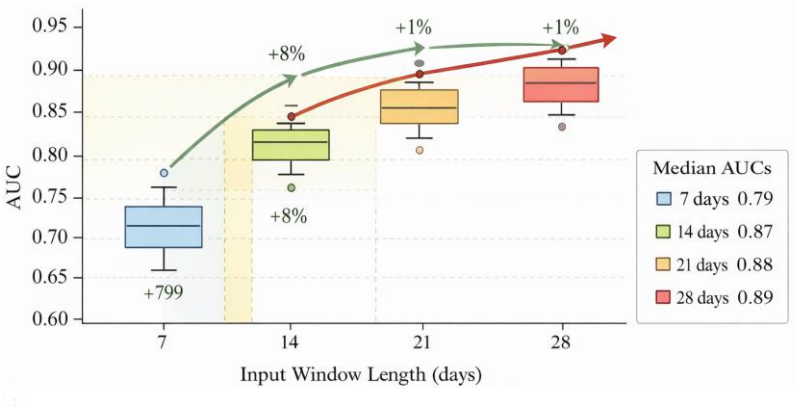

Figure 3. Cross-validation distribution of AUC for the ensemble model under different input windows.

Figure 3 demonstrates that extending the observation window from 7 to 14 days resulted in a substantial increase in mean AUC from 0.79 to 0.87. At the same time, the interquartile range became noticeably narrower, indicating not only improved discriminative capability but also enhanced model stability. This finding suggests that a moderately longer observation period allows the model to capture more comprehensive temporal patterns in patient status changes, thereby improving prediction reliability. When the window was further extended to 14–21 days and 28 days, performance gains became marginal, with median AUC values of 0.88 and 0.89 respectively. This plateau indicates that the additional temporal information beyond two weeks contributes relatively little to further model improvement.

Further examination of the recall distribution revealed that the median recall for the 14-day window reached 0.81, representing an increase of 12 percentage points compared with the 7-day window, while exceeding the 21-day window by only 2 percentage points. These results indicate that the 14-day window achieves an effective balance between detection capability and timeliness. Considering both AUC stability and recall performance, the 14-day window provides sufficient predictive accuracy while maintaining the rapid response required for early screening. Therefore, it can be regarded as an optimal compromise between predictive performance and practical deployment in early clinical risk identification scenarios.

4.4. Importance and Interaction Analysis of Digital Behavioral Features

The SHAP (Shapley Average Absolute Prizes) ranking based on the GBT model showed that the first three non-biological contributors were sleep efficiency (SHAP = 0.21), night activity ratio (SHAP = 0.18) and subjective hunger score (SHAP = 0.15). Together, these three characteristics made up 47% of the overall SHAP, and further two-way interaction testing showed that there was a significant synergy between sleep performance and drug check-in rate (interaction SHAP = 0.06). Subjects with high sleep efficiency (> 85%) and high adherence (clearance rate > 80%) had a 22% higher rate of response than those who had a higher score on either of the individual characteristics alone. The percentage of evening activity was negatively correlated with postprandial hunger scores (interaction SHAP = -0.04). The marginal effect of postprandial hunger on treatment effectiveness was reduced by about 30% when evening activity was higher than the median (32%), indicating that exercise reduced the negative impact of hunger perception on weight loss.Additionally, the interaction term between emotional fluctuation (Likert Scale Standard Deviation) and sleep consistency (WSR) contributed 0.03, suggesting that behavioral rhythm stability enhances the moderating effect of emotion on treatment efficacy. The results show that the main effects and synergistic mechanisms of core behavioral parameters in the low burden digital phenotype can be used to support the selection of individualized intervention objectives.

 doi:10.20944/preprints202603.1764.v1

4.5. Comparative Analysis with Traditional Screening Methods

In order to comprehensively assess the screening effectiveness of the Low Burden Digital Phenotype Framework, it was compared with three conventional screening strategies, and the results are presented in Figure 4.

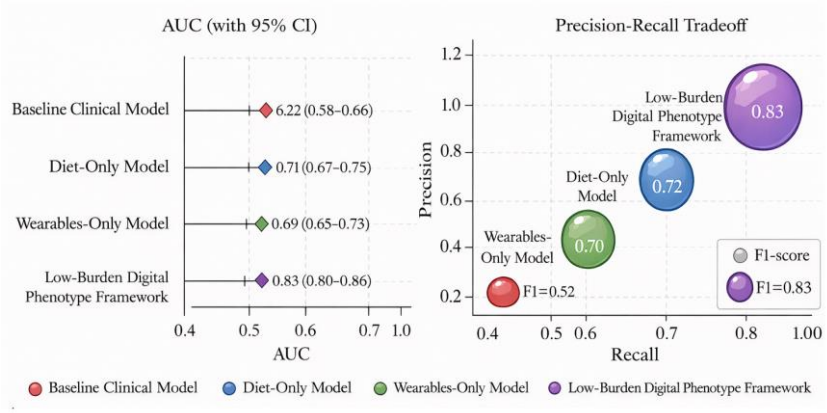

Figure 4. Performance Comparison Between the Low-Burden Digital Phenotype Framework and Traditional Screening Methods.

In Figure 4, a multi-panel forest plot is used to show the key performance metrics of the four strategies over a 14 day window: the left panel shows the main effects, the vertical axis shows the four methods, and the horizontal axis shows the AUC (0.4 - 1.0) .Each method is represented by a point estimation (diamond) and a horizontal line representing 95% confidence. The baseline (age, sex, BMI only) yields 0.62 (0.58 – 0.66), the diet-recording-only model yields 0.71 (0.67 – 0.75), the standalone wearable model yields 0.69 (0.65 – 0.73), while the low-burden digital phenotyping framework achieves 0.83 (0.80 – 0.86). The latter has no overlap with the preceding three, indicating a statistically significant improvement ($p < 0.001$). The right panel shows the two dimensional accuracy and recall distributions at the same time: The size represents the F1-score, with the x-axis showing recall (0.4 – 1.0) and the y-axis showing precision (0.4 – 1.0). Among the four strategies, the low-load digital phenotyping framework is located in the upper right quadrant (recall 0.82, accuracy 0.84) and the largest volume (F1 = 0.83). The primary clinical model (recall 0.55, accuracy 0.58) occupies the lowest left quadrant (F1 = 0.52), whereas the diet and wearable models cluster centrally.Both panels are uniformly color-coded for the four strategies (red, blue, green, violet), with color-strategy maps marked underneath. Results show that the low burden digital phenotyping framework is superior to traditional approaches in terms of AUC, precision recall, and total F1-SCORE, which confirms the incremental value of multimodal, low-burden digital phenotyping for early screening of AOMs.

5. Conclusion

Low burden digital phenotyping has shown great potential for accurate screening of the efficacy of oral anti-obesity drugs. By integrating wearable behavioral data with brief self-report measures, it constructs a "behavioral rhythm – metabolic risk" feature set, achieving a fusion prediction framework that combines static a priori information with dynamic temporal correction.The model delivers high precision response probabilities within a short time window, balancing the need for early screening with the need for application retention. Explanation analysis further reveals the influence of key behavioral indicators — sleep efficiency, night activity, and postprandial hunger perception — on efficacy, providing quantitative evidence for individualized interventions and stratified follow up. Although data collection is limited by light and minimally invasive methods, long term compliance fluctuations and unexplained biological factors may remain. Future improvements can improve the robustness of prediction and clinical scalability through multi-center

 doi:10.20944/preprints202603.1764.v1

validation, enhanced biomarker integration, and optimized prediction window strategies, which provide workable pathways for accurate screening of AOMs and digital health management.

References

1. Tillman E M, Mertami S. Precision medicine to identify, prevent, and treat pediatric obesity[J]. Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy, 2024, 44(12): 939-947.
2. Anazco D, Acosta A. Precision medicine for obesity: current evidence and insights for personalization of obesity pharmacotherapy[J]. International Journal of Obesity, 2025, 49(3): 452-463.
3. Woolf E K, Diktas H E, Acosta A, et al. Precision Prevention, Diagnostics, and Treatment of Obesity: Pipedream or Reality?[J]. Obesity, 2025, 33(11): 2029-2045.
4. Rongala G, Rongala D S, Rongala A N. The Future of Precision Medicine: Targeted Therapies, Personalized Medicine and Formulation Strategies[J]. Journal of Pharmaceutical and BioTech Industry, 2025, 2(4): 19.
5. Uti D E, Alum E U, Atangwho I J, et al. Lipid-based nano-carriers for the delivery of anti-obesity natural compounds: advances in targeted delivery and precision therapeutics[J]. Journal of Nanobiotechnology, 2025, 23(1): 336.
6. Ullah M I, Tamanna S. Obesity: clinical impact, pathophysiology, complications, and modern innovations in therapeutic strategies[J]. Medicines, 2025, 12(3): 19.
7. Drucker D J. Efficacy and safety of GLP-1 medicines for type 2 diabetes and obesity[J]. Diabetes care, 2024, 47(11): 1873-1888.
8. Culver A, Stayrook K, Comerota M, et al. Animal models for development of anti-obesity drugs in the age of GLP-1 agents[J]. Expert Opinion on Drug Discovery, 2025, 20(7): 913-925.
9. Lysaght J, Conroy M J. The multifactorial effect of obesity on the effectiveness and outcomes of cancer therapies[J]. Nature Reviews Endocrinology, 2024, 20(12): 701-714.
10. Yang, Y., Wang, R., Liu, X., Krishnan, A., Tao, Y., Deng, Y., ... & Kong, C. (2025). Declarative Data Pipeline for Large Scale ML Services. arXiv preprint arXiv:2508.15105.