Paper AS22

From Data to Digital Phenotypes: Modelling High-Dimensional PROs and Actigraphy Data in Decentralized Psychiatric Trials

Sakshi Kaushik, ICON, Bangalore, India

ABSTRACT

In recent times, the use of wearable devices and electronic patient reported outcomes (ePRO) has increased rapidly in decentralized psychiatric trials. These trials generate vast amount of high-dimensional longitudinal data. However, the analysis of this type of data poses several statistical challenges such as multicollinearity, time-related dependencies, sparse features, and heterogeneity in patient response. Since it is quite difficult to model such complex data using traditional statistical methods, we introduce a hybrid statistical framework in this study that combines adaptive least absolute shrinkage and selection operator (LASSO) for selecting relevant features and mixed-effects logistic regression for predicting outcomes. This approach handles both within-subject correlation and between-subject variability, providing a new way to analyse high-dimensional data. We tested this framework using simulated data consisting of large number of predictors including actigraphy and PRO measured across several time points for a mental health study. This approach yields interpretable digital phenotypes and informative individual-level predictions.

INTRODUCTION

Decentralized and hybrid clinical trials have become a mainstay in psychiatric research, enabling the passive and active collection of data from wearable devices and self-reported measures (Izmailova et al. 2018; Ticona et al., 2023). The growing adoption of wearable devices and ePRO tools in psychiatric research has made it possible to continuously monitor behavioural and symptom-related data (such as sleep, activity, mood, etc.) outside traditional clinical settings. Such technologies offer the potential for early detection of changes in mental health status, personalization of interventions, and improved trial efficiency. However, the analytical challenges associated with this data are substantial. PRO questionnaires often include dozens of items collected repeatedly over time, and actigraphy can generate a large number of derived features at each assessment. When combined, the resulting datasets are high-dimensional, correlated, and longitudinal in nature. Standard regression approaches are not well suited to this setting, particularly when the number of candidate predictors is large relative to the number of subjects and when within-subject correlations must be accounted for.

Longitudinal PROs are traditionally analysed using multilevel models or growth curve modelling, accounting for repeated measures within subjects. Li and Houts (2021) emphasized the utility of multilevel and multidimensional item response theory (IRT) frameworks for analysing longitudinal PROs in clinical trials, enabling the extraction of latent features (e.g., depression severity) from item-level data while modelling subject-specific trajectories.

The integration of wearable-derived actigraphy data—including sleep patterns, physical activity, and circadian rhythms, etc.—has enhanced the objective quantification of psychiatric symptoms. Studies using data from the RADAR-CNS (Matcham et al., 2019) and UK Biobank cohorts have linked digital endpoints to clinical states in depression and bipolar disorder. Lei et al. (2024) proposed a machine learning-based framework for composite digital biomarkers, employing penalized generalized estimating equations (PGEE) to model longitudinal sensor data.

Recent methodological developments have highlighted the use of penalized regression methods, such as LASSO, for handling high-dimensional data through automated feature selection. The LASSO and adaptive LASSO (Zou, 2006) are commonly used for feature selection in high-dimensional data. Separately, mixed-effects models remain a cornerstone of longitudinal analysis in clinical trials due to their ability to accommodate repeated measurements and subject-level heterogeneity.

Existing statistical approaches tend to focus on either PROs or digital signals independently and often rely on summary measures rather than full-resolution longitudinal data. The statistical methods for jointly analysing longitudinal PRO data and wearable-derived digital endpoints—particularly in high-dimensional settings—remains underdeveloped. Moreover, few models provide both variable selection and the ability to account for within-subject correlations and

population heterogeneity. To address these limitations, we propose a hybrid framework that integrates adaptive LASSO and mixed-effects logistic regression to identify predictive variables and model the binary clinical outcome over time, accounting for repeated measures and random subject effects. The goal is not to introduce a new statistical method, but to demonstrate how existing tools can be used together to support the analysis of high-dimensional PRO and actigraphy data in psychiatric trials. The framework is designed to be transparent, interpretable, and compatible with standard statistical analysis plan (SAP).

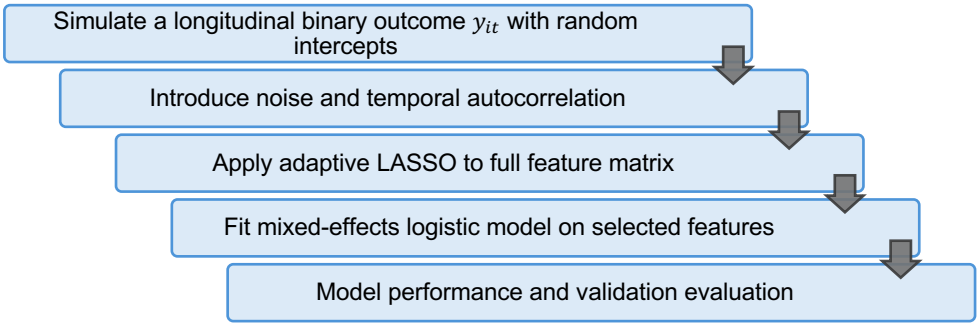
This framework is tested by simulating a psychiatric clinical trial mimicking a remote or hybrid depression study, incorporating 100 digital features and 50 PRO items measured weekly over five time points. The simulation introduces treatment response heterogeneity, temporal correlation, and measurement noise to approximate real-world conditions. The adaptive LASSO regularization method is applied to select relevant digital and PRO predictors and mixed-effects logistic regression is used to account for within-subject correlations. Model performance is assessed using receiver operating curve (ROC) analysis, area under the curve (AUC), and cross-validation. The proposed framework is intended as a secondary or exploratory analysis strategy and as a template that may inform future protocol and SAP development as digital endpoints become more prevalent in clinical research.

STATISTICAL METHODS

This study proposes a two-stage statistical framework for analyzing longitudinal high-dimensional PROs and actigraphy-derived digital endpoints in psychiatric clinical trials. The framework is designed as a secondary and exploratory analytical approach, intended for hypothesis generation and methodological illustration rather than confirmatory inference.

The schema of the statistical methodology implemented is depicted in Figure 1.

Figure 1 Statistical Methodology Schema



NOTATIONS AND DATA STRUCTURE

Assume N subjects are followed over T time points, and each time point includes:

- p PRO items (e.g., fatigue, sleep quality, pain)
- q actigraphy metrics (sleep, activity),

resulting in $p + q$ predictors per subject.

Let: $y_{it} \in \{0, 1\}$: denote a binary clinical improvement indicator variable for subject i at time t $(i = 1, \dots, N; t = 1, \dots, T)$

x_{it}: denote high-dimensional covariate matrix of size $(p + q)$, concatenating PRO and actigraphy features.

Step 1: Adaptive LASSO for feature selection

Firstly, the dimensionality is reduced by applying adaptive LASSO to a logistic regression model:

$$logit\ (\Pr\ (y_{it} = 1)) = x_{it}^{\intercal}\beta$$

With a weighted penalty l_1 to enforce sparsity:

$$\hat{\beta} = arg\min_{\beta}\left\{-l(\beta) + \lambda\sum_{j=1}^{p+q} w_j|\beta_j|\right\}$$

where, $w_j = \frac{1}{|\hat{\beta}_j|^{\gamma}}$ (Zou, 2006; Li et. al., 2017)

Step 2: Mixed-Effects Logistic Regression

Selected predictors $z_{it} \subset x_{it}$ are analysed using below model:

$logit\ (\Pr\ (y_{it} = 1)) = \alpha_i + z_{it}^{\intercal}\beta$

where,

$\alpha_i \sim N(0, \sigma^2)$ are subject-level random effect (Verbeke & Molenberghs, 2000; Fitzmaurice et. al., 2011).

Step 3: Model Accuracy

Model Accuracy is evaluated using multiple complementary metrics (Steyerberg et al, 2019)

- AUC and ROC curves for discrimination
- Classification accuracy at a 0.5 threshold
- Mean squared error (MSE) and root mean squared error (RMSE) for LASSO model
- Brier score for probabilistic accuracy for mixed-effects logistic regression model
- Calibration plots comparing predicted probabilities with observed response rates

These diagnostics are intended to demonstrate model adequacy and stability and not comparisons.

All analyses were conducting using R (version 4.5.2). The complete simulation and analysis code used in this study is provided in the Appendix to support transparency and reproducibility.

RESULTS

SIMULATION STUDY

We simulated a decentralized psychiatric trial with:

- 200 subjects measured weekly over 5 weeks
- 150 total predictors with 50 PRO items (mood, fatigue, concentration, anxiety, insight) and 100 actigraphy endpoints (e.g., sleep efficiency, movement, circadian metrics)
- The outcome variable, a binary depression response which is a marker of clinical improvement with 1 = symptom improvement and 0 = No improvement (e.g., ≥ 50% reduction in depression score) was generated using a logistic model with prespecified subset of 10 truly predictive features. Subject specific random intercepts were included to induce within-subject correlation across repeated measurements.
- The simulation was designed to reflect realistic challenges encountered in digital biomarker-enabled trial, including high dimensionality, correlated predictors and longitudinal dependence.

The following are the first few rows and columns of data frame consisting of 1000 observations on 153 variables generated for 200 subjects.

subject_id	Time	Y	X1	X2	X3	X4	X5	X6	X7
SUBJ-001	1	0	-0.4120734	0.9115762	0.2416976	0.3539397	1.6957224	0.2946438	-0.5540833
SUBJ-001	2	1	0.2941283	-0.3296327	0.4324846	-1.0738541	-1.331133	-0.4796764	0.5140281
SUBJ-001	3	1	-0.8134725	-1.9973868	-0.0076093	0.400616	0.8276127	1.0543189	-0.3780314
SUBJ-001	4	0	0.4128176	-2.2617034	-0.6130389	0.0982705	1.3124019	0.0669615	-1.0074358
SUBJ-001	5	1	1.7753429	1.0556374	-0.0177988	0.1720391	0.9634209	0.9317492	1.088856
SUBJ-002	1	0	-0.2822167	1.1762367	0.2955424	-1.1053203	-1.1439064	1.330151	0.4292197
SUBJ-002	2	1	1.8067315	-1.021326	0.7428111	0.0601259	0.6833505	-1.3082496	-0.4713088
SUBJ-002	3	0	-0.0984981	-0.2731733	-1.0392731	-1.3607781	-0.5737859	0.2780263	2.1199003
SUBJ-002	4	0	-1.8778929	1.5439964	-0.1301572	-0.3835864	-1.0917345	-0.141197	1.312886

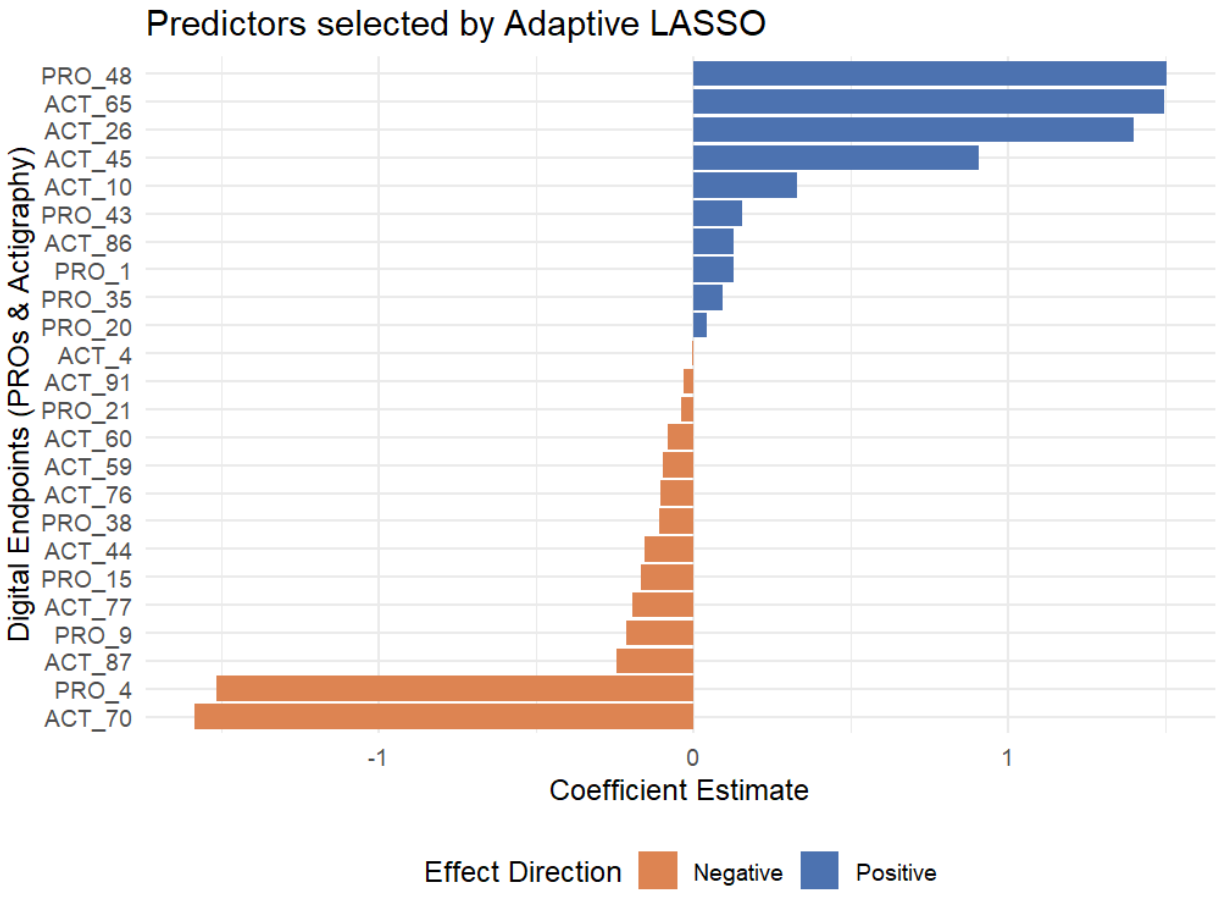
SUBJ-002	5	1	-0.2115311	0.2935456	-1.1142595	-0.7213061	-0.3595806	0.5788807	0.5558715
SUBJ-003	1	1	-1.027221	0.9340997	0.9306266	-0.9369478	-1.4129139	-0.9280003	1.2633536
SUBJ-003	2	1	1.7599568	-1.6956803	-1.0759263	-0.0868296	-1.4445268	0.0030897	0.9632685
SUBJ-003	3	0	0.7077767	1.1004915	-0.3614914	0.4311248	2.0085437	0.5074465	-0.3011768
SUBJ-003	4	1	-0.4187638	-0.5987606	0.1444684	1.0173233	-0.6925844	0.4395817	-0.5907733
SUBJ-003	5	1	-0.9703147	1.763387	1.1565994	2.0663701	0.0906722	1.2577915	1.2575704

FEATURE SELECTION VIA ADAPTIVE LASSO

The adaptive LASSO penalty parameter was selected using 10-fold cross validation with the optimal value $\lambda = 0.5268$, chosen at the minimum cross-validated deviance. The adaptive LASSO identified 24 predictors out of 150 features including 10 PRO and 14 actigraphy-derived variables. Although simulation incorporated 10 truly predictive features, additional correlated predictors were retained, reflecting redundancy and shared information commonly observed in digital and PRO data.

Figure 2 presents the coefficient plot for predictors retained by the adaptive LASSO, with coefficients color-coded to distinguish positive and negative associations with the binary treatment response. Predictors with positive coefficients indicate increased odds of response, whereas negative coefficients correspond to reduced odds of response. The magnitude of coefficient estimates varied across predictors, suggesting that predictive signal is distributed across multiple features rather than dominated by a single variable.

Figure 2 Coefficient plot of adaptive LASSO-selected predictors, with colours indicating direction of association

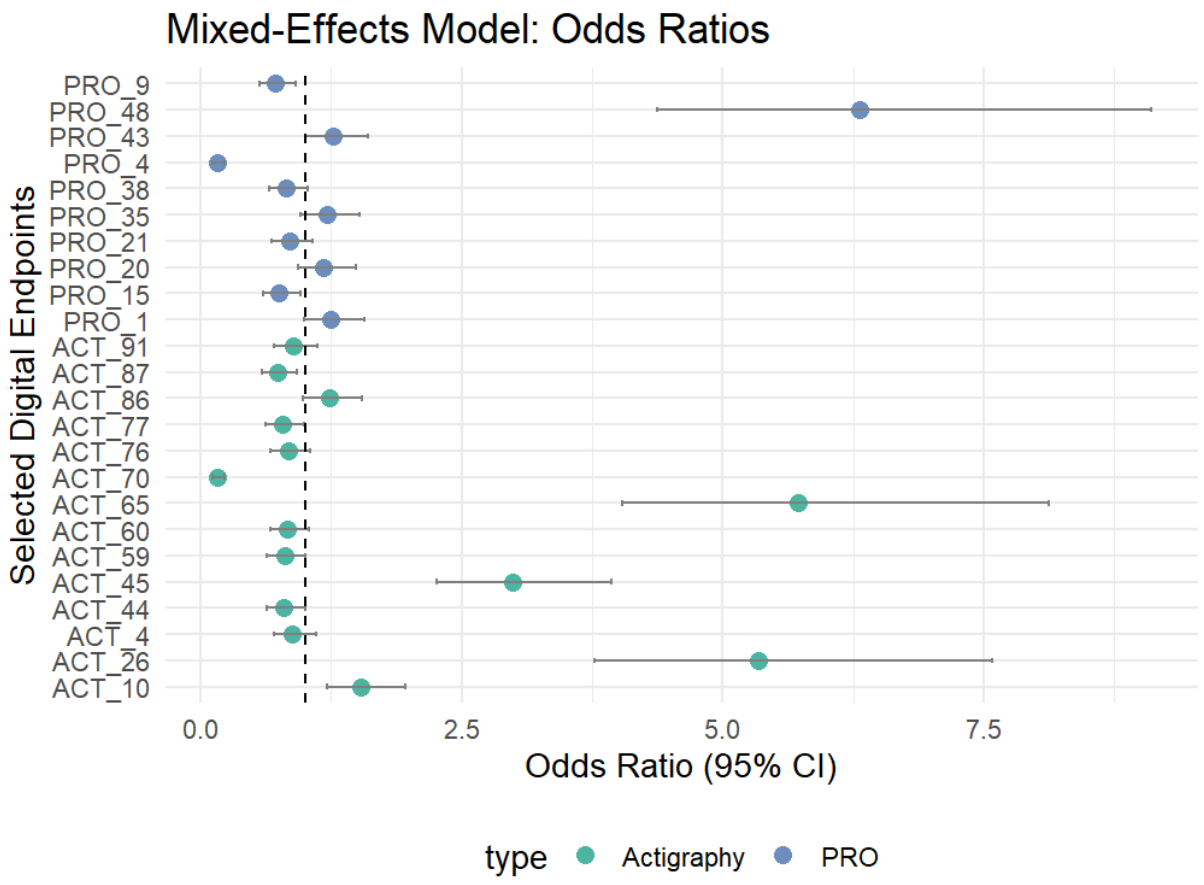


MIXED-EFFECTS LOGISTIC REGRESSION RESULTS

Predictors retained by adaptive LASSO were subsequently entered into mixed-effects logistic regression model to evaluate longitudinal effects while accounting for within-subject correlation arising from repeated measurements over time. The model converged successfully using maximum likelihood estimation.

Estimated fixed effects from the mixed-effects model were summarized using odds ratios (ORs) and corresponding 95% confidence intervals. Figure 3 presents a forest plot of ORs for the selected predictors. Predictors with ORs greater than one indicate increased odds of treatment response, whereas ORs less than one indicate reduced odds. The direction and magnitude of effects varied across predictors, reflecting heterogeneous contributions of PROs and actigraphy features to treatment response.

Figure 3 Forest plot of odds ratios and 95% CIs for fixed effects from mixed-effects logistic regression model



MODEL PERFORMANCE

MODEL PERFORMANCE OF LASSO MODEL

Although adaptive LASSO was primarily used for feature screening rather than inference, descriptive performance metrics were computed to assess whether the retained features preserved meaningful predictive information. The adaptive LASSO model achieved AUC value of 0.9416 indicating acceptable discrimination between responders and non-responders based on the high-dimensional PRO and actigraphy data. Figure 4 presents the ROC curve illustrating discriminative performance of adaptive LASSO model. Classification accuracy was evaluated at a probability threshold of 0.5, yielding an accuracy of 0.859 indicating a satisfactory model performance. Prediction error was quantified using MSE and RMSE, calculated from predicted probabilities. The adaptive LASSO yielded MSE of 0.098 and RMSE of 0.313, reflecting moderate deviation between predicted probabilities and observed outcomes. Collectively, these results

indicate that the selected predictors collectively retained the predictive signal, defined here as the ability to discriminate the responders from non-responders. This supports that the selected predictors are suitable for longitudinal modelling.

MODEL PERFORMANCE OF MIXED-EFFECTS LOGISTIC REGRESSION MODEL

The mixed-effects logistic regression model achieved an AUC value of 0.9435 (being close to 1) indicating that the model has performed well in discriminating between positive and negative classes. Figure 4 presents the ROC curve illustrating discriminative performance of mixed-effects logistic regression model. Using a probability threshold of 0.5, the model yielded a classification accuracy of 0.869 indicating that the model's predictions matched the actual outcomes in 86.9% of the cases. These results indicate that the model was able to distinguish responders from non-responders while accounting for longitudinal correlation. Probabilistic accuracy assessed using the Brier score yielded a value of 0.0963, indicating reasonable agreement between predicted probabilities and observed outcomes. Together, these results validate the adequacy of the mixed-effects model, demonstrating its ability to capture longitudinal response patterns without inflating predictive performance metrics. Calibration of the mixed-effects model was evaluated by comparing predicted probabilities with observed response rates. As shown in

Figure 5, predicted probabilities aligned reasonably well with observed proportions, suggesting satisfactory calibration at the population level. Collectively, these results demonstrate that the mixed-effects logistic regression model is suitable to perform longitudinal analysis of high dimensional digital and PRO data.

Figure 4 ROC curves illustrating discriminative performance of adaptive LASSO and mixed-effects regression model

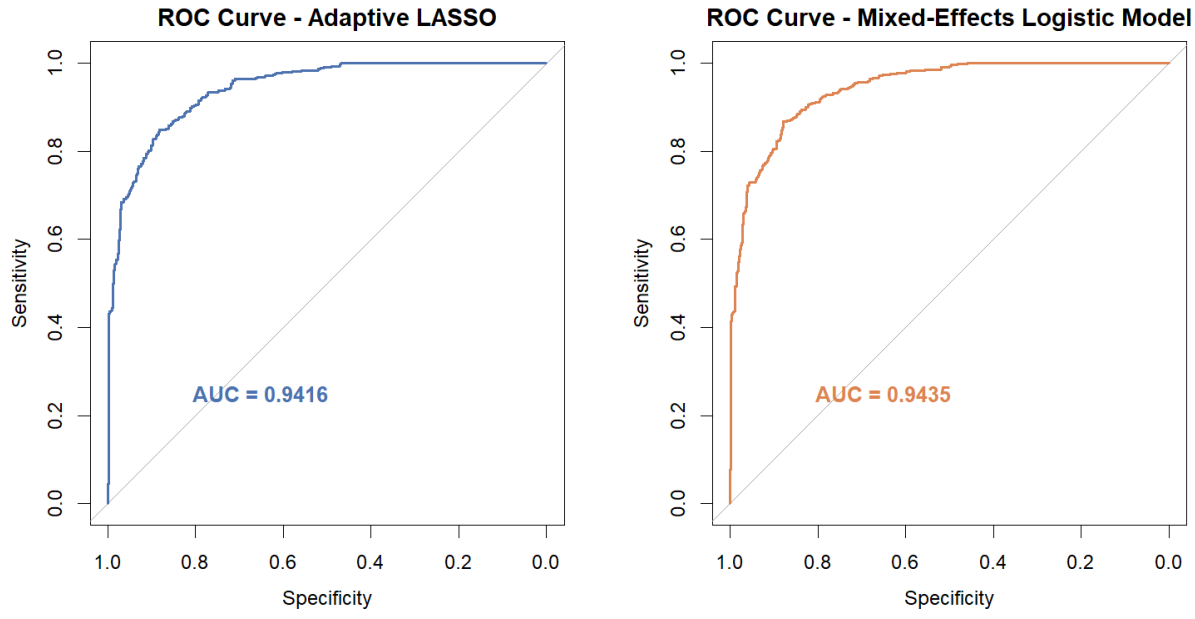
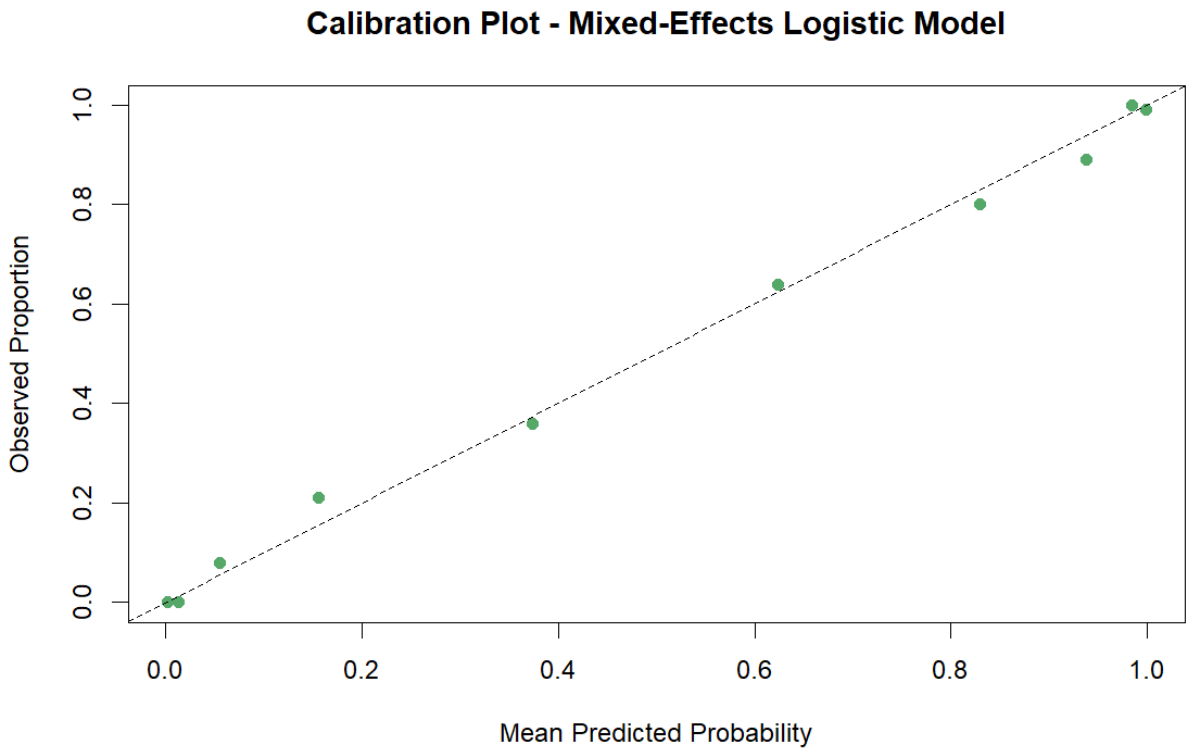


Figure 5 Calibration plot comparing predicted probabilities and observed response rates for mixed-effects logistic regression model



DISCUSSION

The increasing use of patient-reported outcomes and digital endpoints such as actigraphy has created new analytical challenges for psychiatric clinical trials. These data are typically high-dimensional, correlated, and collected repeatedly over time, making conventional modeling approaches difficult to apply without substantial simplification. This study presents a practical way to address these challenges by combining adaptive LASSO feature screening with mixed-effects logistic regression in a structured and transparent workflow.

In this study, adaptive LASSO was used as an exploratory screening step to reduce the dimensionality of a large set of PRO and actigraphy variables. In the simulated setting, the method selected a subset of predictors that exceeded the number of truly simulated features, which is expected in correlated and noisy high-dimensional data. Rather than viewing this as a limitation, this behavior reflects the sensitivity of adaptive LASSO to select weak but potentially relevant predictors and supports its use for hypothesis generation rather than confirmatory inference.

The second stage of the framework applied mixed-effects logistic regression to the selected predictors, allowing longitudinal correlations and between-subject heterogeneity to be explicitly modeled. This step enabled clinically

interpretable inference through odds ratios while preserving the repeated-measures structure of the data. The estimated random intercept variance highlighted meaningful differences in baseline response rates across subjects, reinforcing the importance of accounting for subject-level effects in psychiatric trials.

Model performance metrics were reported for both stages to provide a context on model adequacy rather than to compare models. The adaptive LASSO demonstrated acceptable discrimination and prediction error, indicating that the screened predictors retained meaningful predictive information. The mixed-effects model showed consistent discrimination, reasonable calibration, and stable prediction error, suggesting that the longitudinal model provided a reliable representation of response patterns over time.

An important aspect of this work is its focus on interpretability and implementation. Visual summaries such as coefficient plots, odds-ratio forest plots, and ROC curves helped translate complex high-dimensional results into outputs that are accessible to statisticians, clinicians, and trial stakeholders. All analyses were conducted using standard R packages that are widely used in clinical trial settings, supporting reproducibility and ease of adoption.

Overall, the value of this study lies not in proposing a new statistical model, but in demonstrating how existing, well-understood methods can be combined thoughtfully to address emerging challenges posed by digital and PRO data. The framework aligns well with current regulatory expectations by clearly separating exploratory feature screening from longitudinal inference and by positioning the analysis as a secondary or hypothesis-generating approach.

CONCLUSION

This paper presents a statistically robust and operationally feasible approach for exploratory analysis of longitudinal high-dimensional PRO and digital endpoint data in psychiatric clinical trials. Although no novel statistical models are introduced, practitioners working with digital biomarkers and decentralized trial data will find significant value in the practical methodology proposed, which integrates existing tools and positions them effectively for use in clinical trial settings. This framework may serve as a template for future protocol pre-specification, particularly in early-phase trials or exploratory objectives where hypothesis generation is a key goal. With the expanding usage of digital endpoints in clinical trials, this framework offers a practical way to support monitoring, improve endpoint definitions, and generate evidence.

LIMITATIONS

We observed some limitations to this study. Firstly, the framework was evaluated using simulated data, which cannot fully capture the complexity, missingness patterns, and behavioral variability observed in real-world psychiatric trials. Secondly, adaptive LASSO tends to select more predictors than truly causal, particularly in the presence of correlated features, which reinforces the importance of treating results as exploratory rather than confirmatory.

Thirdly, formal cross-validation of the mixed-effects logistic regression model was not implemented, as standard cross-validation procedures are less straightforward for hierarchical models. Instead, model adequacy was assessed using discrimination, calibration, and error-based metrics, which provide complementary but not exhaustive validation.

Finally, this work focuses on binary outcomes; extensions to continuous or time-to-event endpoints were not explored.

FUTURE SCOPE

We propose few extensions to this framework for future study. To ensure reliability, the approach should be applied to actual psychiatric clinical trial data to evaluate its performance under realistic conditions, including missing data and inconsistent patient adherence.

Methodologically, future work may explore stability selection, group LASSO, or Bayesian hierarchical shrinkage approaches as alternatives or complements to adaptive LASSO. Time-varying effects could also be incorporated in the modelling.

REFERENCES

1. Izmailova Es, Wagner JA, Perakslis ED (2018). Wearable devices in clinical trials: hype and hypothesis. Nat Rev Drug Discov.

2. Ticona, E et al. (2023). Decentralised, patient-centric, site-less, virtual and digital clinical trials? Drug Discovery today, 28, 103520.
3. Matcham F et al. (2019). Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD). JMIR
4. Zou, H. (2006). The Adaptive Lasso and Its Oracle Properties. Journal of the American Statistical Association.
5. Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software.
6. Lei, Y., et al. (2024). A Machine Learning Framework for Digital Biomarkers. JMIR.
7. Li, C., & Houts, C. (2021). Longitudinal Analysis of PROs in Clinical Trials. Journal of Biopharmaceutical Statistics.
8. Steyerberg et al (2019). Clinical prediction models. Springer.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Author Name: Sakshi Kaushik, Ph.D

Company: ICON

Email: Sakshi.kaushik@iconplc.com

APPENDIX : R CODE FOR DATA SIMULATION AND STATISTICAL ANALYSIS

This appendix provides the complete R code used to implement the simulation study and exploratory statistical analyses described in above sections. The code demonstrates the generation of longitudinal high-dimensional data, including PROs and actigraphy-derived digital endpoints, and the application of adaptive LASSO for feature selection followed by mixed-effects logistic regression. The code is extended to include the assessment of model performance and various visualizations.

All analyses were conducted using R (version 4.5.2) and established packages, including *glmnet* for penalized regression and *lme4* for mixed-effects modelling. The scripts are provided to support transparency and reproducibility and are intended for illustrative and education purposes.

```
# Required packages
library(MASS)
library(glmnet)
library(lme4)
library(caret)
library(pROC)
library(ggplot2)
library(broom.mixed)
library(dplyr)

set.seed(13254)

# ------------------------------
# 1. Simulate Longitudinal Data
# ------------------------------

n <- 200   # subjects
t <- 5     # time points
p_pro <- 50 # 50 PRO features
p_act <- 100 # 100 actigraphy features
p <- p_pro + p_act # total features
true_features <- 10  # true predictors

# Create subject IDs and time
```

```r
# Generate sequential IDs with leading zeros (e.g., 001, 002)
prefix <- "SUBJ-"
id <- sprintf("%s%03d", prefix, 1:n)
subject_id <- rep(id, each = t)
Time <- rep(1:t, times = n)


# Simulate random effects (subject-level intercepts)
rand_int <- rnorm(n, 0, 1)
rand_int_rep <- rep(rand_int, each = t)


# Simulate feature matrix
# Simulate PRO features
X_pro <- matrix(rnorm(n * t * p_pro), nrow = n * t, ncol = p_pro)

# Simulate actigraphy features
X_act <- matrix(rnorm(n * t * p_act), nrow = n * t, ncol = p_act)

# Combine: PRO first and then actigraphy
X <- cbind(X_pro, X_act)
colnames(X) <- c(
  paste0("PRO_", 1:p_pro),
  paste0("ACT_", 1:p_act)
)

# Assign true beta coefficients for 10 features
beta <- rep(0, p)
beta[sample(1:p, true_features)] <- runif(true_features, -2, 2)


# Linear predictor with subject random effect
eta <- X %*% beta + rand_int_rep


# Convert to probabilities and simulate binary outcome
prob <- 1 / (1 + exp(-eta))
Y <- rbinom(n * t, 1, prob)


# Put into data frame
data <- data.frame(subject_id = as.factor(subject_id), Time = Time, Y = Y, X)


# ----------------------------------------
# 2. Adaptive LASSO for variable selection
# ----------------------------------------

# First, fit regular logistic regression to get initial beta estimates
cvfit_initial <- cv.glmnet(X, Y, family = "binomial", alpha = 1)
beta_initial <- as.vector(coef(cvfit_initial, s = "lambda.min"))[-1]

# Adaptive weights
gamma <- 1
w <- 1 / (abs(beta_initial)^gamma + 1e-4)  # add small constant to avoid div-
by-zero
```

```r
# Fit adaptive LASSO
X_scaled <- scale(X)
cvfit_adalasso <- cv.glmnet(X_scaled, Y, family = "binomial", alpha = 1,
penalty.factor = w)

# Optimal lambda using the minimum criterion
optimal_lambda_min <- round(cvfit_adalasso$lambda.min, 4)

# Extract coefficient estimates from the adaptive LASSO model
# using the optimal penalty parameter
lasso_coef <- coef(cvfit_adalasso, s="lambda.min")

# Convert sparse coefficient matrix to a data frame
lasso_coef_df <- data.frame(
  feature = rownames(lasso_coef),
  coefficient = as.numeric(lasso_coef))

# Retain only predictors with non-zero coefficients
selected_lasso<-subset(
  lasso_coef_df,
  coefficient!=0 & feature!="(Intercept)"
)

nrow(selected_lasso) # number of predictors selected

pro_selected <- sum(grepl("^PRO_", selected_lasso)) # no. of PRO predictors
selected
act_selected <- sum(grepl("^ACT_", selected_lasso)) # no. of ACT predictors
selected

# Visualization: Adaptive LASSO selected predictors

ggplot(selected_lasso,
       aes(x=reorder(feature, coefficient), y=coefficient,
           fill=coefficient > 0)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  scale_fill_manual(
    values = c("#DD8452", "#4C72B0"),
    labels = c("Negative", "Positive"),
    name = "Effect Direction"
  ) +
  labs(
    title="Predictors selected by Adaptive LASSO",
    x="Digital Endpoints (PROs & Actigraphy)",
    y="Coefficient Estimate"
  ) +
  theme_minimal(base_size = 12) +
  theme(legend.position = "bottom")

# ------------------------------------
# 3. Mixed-Effects Logistic Regression
# ------------------------------------

# Create modeling dataset
```

```r
#selected_names <- paste0("X", selected_vars)
lasso_coef <- coef(cvfit_adalasso, s="lambda.min")

selected_vars <- rownames(lasso_coef)[lasso_coef[,1] != 0]
selected_vars <- selected_vars [selected_vars != "(Intercept)"]
model_data <- data[, c("subject_id", "Y",selected_vars)]


# Fit model
formula <- as.formula(paste("Y ~", paste(selected_vars, collapse = " + "), "+
(1 | subject_id)"))
mixed_model <- glmer(formula,
  data = model_data,

  family = binomial,
  control = glmerControl(optimizer = "Nelder_Mead", optCtrl = list(maxfun =
100000)))

summary(mixed_model)

# Odds Ratio + CI

or_df <- tidy(
  mixed_model,
  effects = "fixed",
  conf.int = TRUE,
  exponentiate = TRUE) # gives Odds Ratio
#or_df <- or_df %>%
 # filter(term != "(Intercept)")
or_df <- or_df[or_df$term != "(Intercept)", ]

# Forest Plot for Odds Ratio
or_df <- or_df %>%
  mutate(type = case_when(
    grepl("^PRO", term) ~ "PRO",
    grepl("^ACT", term) ~ "Actigraphy"
  ))

ggplot(or_df, aes(x=estimate, y=term, color = type)) +
  geom_point(size = 3) +
  geom_errorbarh(aes(xmin = conf.low, xmax = conf.high), height = 0.2, color
= "grey50") +
  geom_vline(xintercept = 1, linetype = "dashed", color = "black") +
  labs(
    title = "Mixed-Effects Model: Odds Ratios",
    x = "Odds Ratio (95% CI)",
    y = "Selected Digital Endpoints"
  ) +
  scale_color_manual(values = c("PRO" = "#6C8EBF", "Actigraphy" = "#4CB5A2"))
+
  theme_minimal(base_size = 14) +
  theme(legend.position = "bottom")



# ------------------------------
```

```r
# 4. Model performance
# ------------------------------

# ---------------------------------
# Model performance for LASSO model
# ---------------------------------
# Predicted probabilities from LASSO model

lasso_coef_df1<-subset(
  lasso_coef_df,
  feature!="(Intercept)"
)

pred_lasso <- predict(
  cvfit_adalasso,
  newx = X,
  s = "lambda.min",
  type = "response"
  )

# ROC and AUC for LASSO
roc_lasso <- roc(Y, as.numeric(pred_lasso))
auc_lasso <- auc(roc_lasso)


# Classification accuracy at 0.5 threshold
pred_class_lasso <- ifelse(pred_lasso > 0.5, 1, 0)
acc_lasso <- mean(pred_class_lasso == Y)

# Mean squared error for LASSO model
mse <- round(mean((pred_lasso - Y)^2),3)
rmse <- round(RMSE(pred_lasso, Y), 3)


# ----------------------------------------
# Model performance for Mixed-Effects model
# ----------------------------------------
# Population-average predicted  probabilities (Fixed Effects only)

pred_mixed <- predict(
  mixed_model,
  newdata = model_data,
  type = "response",
  re.form = NA) # excludes random effects

# ROC and AUC for LASSO
roc_mixed <- roc(model_data$Y, as.numeric(pred_mixed))
auc_mixed <- auc(roc_mixed)

# Classification accuracy at 0.5 threshold
pred_class_mixed <- ifelse(pred_mixed > 0.5, 1, 0)
acc_mixed <- mean(pred_class_mixed == model_data$Y)

# Visualization: ROC plot for LASSO and Mixed-Effects models

par(mfrow = c(1, 2))
plot(roc_lasso,
```

```
     main = "ROC Curve - Adaptive LASSO",
     col = "#4C72B0",
     lwd = 2)

# Add AUC text inside the LASSO ROC plot
text(
  x = 0.65,
  y = 0.25,
  labels = paste0("AUC = ", round(auc_lasso, 4)),
  col = "#4C72B0",
  cex = 1.1,
  font = 2
)

plot(roc_mixed,
     main = "ROC Curve - Mixed-Effects Logistic Model",
     col = "#DD8452",
     lwd = 2)

# Add AUC text inside the Mixed-Effects ROC plot
text(
  x = 0.65,
  y = 0.25,
  labels = paste0("AUC = ", round(auc_mixed, 4)),
  col = "#DD8452",
  cex = 1.1,
  font = 2
)
par(mfrow = c(1, 1))

# Brier Score for Mixed-Effects logistic model
Br_scr <- round(mean((pred_mixed - model_data$Y)^2),4)

#Calibration plot for Mixed-Effects model
calib_df <- data.frame(
  obs = model_data$Y,
  pred = pred_mixed
) %>%
  mutate(bin = ntile(pred, 10)) %>%
  group_by(bin) %>%
  summarise(
    mean_pred = mean(pred),
    mean_obs = mean(obs)
  )

plot(calib_df$mean_pred,calib_df$mean_obs,
     xlab = "Mean Predicted Probability",
     ylab = "Observed Proportion",
     main = "Calibration Plot - Mixed-Effects Logistic Model",
     pch = 19, col = "#55A868"
     )
abline(0, 1, lty = 2)
```