

# Longitudinal Modeling of 30-Day Hospital Readmission Among Adults With Diabetes

## AUTHORS

Zhangziyan Jiang

Wanbing Wang

Xiaochen Zhou

Muyao Tang

Danni Liu

## Introduction

Hospital readmissions among adults with diabetes impose a substantial burden on patients and health systems, and 30-day readmission has become a widely used quality metric. Prior work has identified demographic factors, comorbidity burden, and prior healthcare utilization as important correlates of early readmission, but far less is known about how risk evolves over repeated encounters for the same individual and whether longitudinal trajectories differ by treatment patterns such as insulin use or intensification. Understanding these dynamics is essential for designing targeted interventions that prioritize patients who remain persistently high-risk despite ongoing contact with the healthcare system.

Leveraging an electronic health record–derived cohort of diabetic inpatients with multiple hospital encounters, we examined the association between insulin regimen, encounter history, and 30-day readmission. We first used a marginal logistic GLM to characterize population-averaged associations and evaluate whether a simple independence model provides an adequate description of the data. We then assessed within-patient correlation and formally compared this marginal approach with subject-specific generalized linear mixed models that incorporate random intercepts and random slopes for encounter index.

Our primary aims were threefold: (i) to quantify how 30-day readmission risk changes with successive encounters, (ii) to determine whether these trajectories differ across insulin treatment groups, particularly patients undergoing insulin dose escalation, and (iii) to evaluate the extent to which accounting for within-patient correlation changes substantive conclusions. By integrating model diagnostics, robust standard errors, and mixed-effects modeling, we

sought to provide a rigorous and clinically interpretable description of longitudinal readmission risk in this high-risk population.

## Data

We analyzed 1488 unique patients contributing 4160 inpatient encounters from an existing de-identified administrative database of adults with diabetes. The overall 30-day readmission rate was 19.4%, indicating that early readmission is a frequent outcome in this population with repeated hospitalizations. Across insulin regimens, the numbers of unique patients ever observed on each regimen, together with the corresponding encounter counts and 30-day readmission proportions, are summarized in the table above. Because some individuals transition between regimens over time, the patient counts by regimen sum to more than the overall number of unique patients.

Sample size and 30-day readmission by insulin regimen

Insulin regimen	Patients (per regimen)	Encounters	30-day readmission
Down	396	591	20.6%
No	908	1745	17.5%
Steady	778	1256	20.9%
Up	369	568	20.8%

## GLM

GLM Method:

Let  $i = 1, \dots, N$  index patients and  $j = 1, \dots, n_i$  index encounters within patient  $i$ . For each encounter we observe a binary outcome

$$Y_{ij} = \begin{cases} 1, & \text{if patient } i \text{ is readmitted within 30 days after encounter } j, \\ 0, & \text{otherwise,} \end{cases}$$

and a vector of covariates  $\mathbf{X}_{ij}$  including insulin regimen, demographics, encounter-level severity indicators, and prior utilization measures. As an initial marginal analysis, we modeled the encounter-level readmission probability

using a logistic generalized linear model (GLM) with a logit link,

$$\text{logit}\{\Pr(Y_{ij} = 1)\} = \beta_0 + \boldsymbol{\beta}_{\text{Ins}}^\top \mathbf{Z}_{\text{Ins},ij} + \boldsymbol{\beta}_{\text{Demo}}^\top \mathbf{Z}_{\text{Demo},ij} + \boldsymbol{\beta}_{\text{Util}}^\top \mathbf{Z}_{\text{Util},ij} + \beta_{\text{Enc}} \text{Time}_{ij},$$

where  $\mathbf{Z}_{\text{Ins},ij}$  encodes insulin regimen (No, Steady, Up, Down),  $\mathbf{Z}_{\text{Demo},ij}$  collects age group, gender, and race,  $\mathbf{Z}_{\text{Util},ij}$  contains encounter-level severity and prior utilization measures (number of medications, length of stay, numbers of prior emergency and inpatient visits, admission type), and  $\text{Time}_{ij}$  denotes the truncated encounter index (capped at 5). This model assumes that, conditional on  $\mathbf{X}_{ij}$ , the  $\{Y_{ij}\}$  are independent across encounters and across patients. Regression coefficients were exponentiated to obtain odds ratios (ORs), so that  $\exp(\beta_k)$  is interpreted as the multiplicative change in the odds of readmission associated with a one-unit increase in the corresponding covariate, holding other variables fixed. Wald  $z$ -tests were used for inference on regression coefficients.

With respect to covariate selection, we adopted a hypothesis-driven strategy rather than data-driven stepwise procedures. Demographic variables (age group, gender, race) and clinical severity or utilization indicators (admission type, length of stay, numbers of prior emergency and inpatient visits, number of medications) were retained regardless of individual  $p$ -values in order to control for potential confounding and to preserve comparability with prior readmission studies. Time and its interaction with insulin regimen were then added explicitly to address our primary scientific question about longitudinal trajectories under different treatment patterns.

To evaluate the contribution of encounter history, we fit two versions of the GLM: one including the truncated encounter index and one excluding it. Comparing these models allowed us to assess whether encounter index added explanatory value and should be retained before proceeding to more advanced longitudinal modeling frameworks.

Regression coefficients were exponentiated to obtain odds ratios (ORs), and profile likelihood confidence intervals were computed using `confint()`. This allowed effects to be interpreted as multiplicative changes in the odds of readmission while maintaining likelihood-based inference.

Generalized variance inflation factors (GVIFs) were calculated to evaluate multicollinearity, particularly for categorical predictors with multiple degrees of freedom. This ensured stable coefficient estimation and interpretability.

We performed residual diagnostics, influence diagnostics, and goodness-of-fit assessments to evaluate whether the GLM adequately captured the relationship between predictors and 30-day readmission. Deviance and Pearson residuals were plotted against fitted probabilities to assess functional form, detect outlying patterns, and evaluate the appropriateness of the logit link. Overdispersion was examined by comparing the residual deviance to its degrees of freedom. Influence diagnostics were conducted using Cook's distance to identify observations with disproportionate leverage on coefficient estimates. Model calibration was assessed using the Hosmer–Lemeshow test, and predictive discrimination was evaluated using the receiver operating characteristic (ROC) curve and its corresponding area under the curve (AUC). Together, these diagnostics provided a comprehensive assessment of model fit and predictive performance.

Because encounters were nested within patients, the independence assumption of a standard GLM may be violated. To explore potential within-patient dependence, we first generated encounter-level trajectory plots, which display how observed readmission proportions vary across encounter index and insulin categories. These plots provide an initial visual assessment of temporal patterns and possible effect modification over repeated encounters.

Next, we computed lag-1 Pearson residuals, plotting residuals at encounter  $t$  against residuals from encounter  $t-1$ . This diagnostic examines serial correlation among consecutive encounters; systematic dependence or diagonal structure in this plot would indicate violation of the independence assumption.

To evaluate whether the association between insulin status and readmission varied across encounter history, we fit a logistic GLM including insulin  $\times$  encounter\_index\_trunc interaction terms. This model tested whether the effect of insulin depended on how many encounters a patient had previously accumulated. Model selection using AIC guided whether the interaction terms improved model specification. Conceptually, one could also entertain richer mean structures that allow for non-linear time effects and additional interactions (e.g., quadratic terms in  $\text{Time}_{ij}$ , and interaction terms between insulin regimen and prior utilization), but given the modest number of encounters per patient and our focus on a single, clinically interpretable interaction, we restricted our implemented analyses to the linear-time specification with an insulin-by-time interaction.

Formally, the interaction model can be written as

$$\text{logit}\{\Pr(Y_{ij} = 1)\} = \beta_0 + \boldsymbol{\beta}_{\text{Ins}}^\top \mathbf{Z}_{\text{Ins},ij} + \beta_{\text{Time}} \text{Time}_{ij} + \boldsymbol{\beta}_{\text{Int}}^\top (\mathbf{Z}_{\text{Ins},ij} \times \text{Time}_{ij}) + \boldsymbol{\gamma}^\top \mathbf{W}_{ij},$$

where  $\mathbf{W}_{ij}$  denotes the remaining covariates (demographics, admission type, prior utilization, and encounter-level

severity). In this parameterization,  $\beta_{\text{Time}}$  captures the time trend in the reference insulin group (No insulin), while each component of  $\beta_{\text{Int}}$  quantifies how the time trend is modified in the corresponding insulin category. Thus  $\exp\{\beta_{\text{Time}} + \beta_{\text{Int},k}\}$  gives the multiplicative change in the odds of readmission per one-unit increase in encounter index for insulin group  $k$ , relative to the baseline group.

Finally, to account for potential correlation directly in the inferential framework, we computed cluster-robust (sandwich) standard errors at the patient level. These robust SEs adjust for within-patient clustering while leaving coefficient estimates unchanged, allowing valid inference even when observations are correlated. Comparing model-based and cluster-robust standard errors helped assess the extent to which within-patient dependence affected uncertainty estimates and whether a marginal GLM remained adequate or if repeated-measures models (GEE or GLMM) might be required.

#### GLM Result:

Here two logistic GLMs were fitted. One includes a truncated encounter index(capped at 5), and one without it. Including encounter\_trunc improved model fit (AIC 4026 vs 4047) and reduced residual deviance (3992.2 vs 4015). The truncated index was highly significant( $p<0.001$ ).

Odds Ratios (OR) with 95% Confidence Intervals for Logistic GLM

	Term	OR	2.5 %	97.5 %	p.value
(Intercept)	Intercept	0.2683	0.1906	0.3758	0.0000
insulinDown	Insulin: Down vs No	1.1268	0.8805	1.4352	0.3379
insulinSteady	Insulin: Steady vs No	1.2006	0.9947	1.4483	0.0564
insulinUp	Insulin: Up vs No	1.1606	0.9031	1.4841	0.2396
age>=80	Age: $\geq 80$ vs $\leq 50$	0.8733	0.6689	1.1414	0.3201
age50-80	Age: 50–80 vs $\leq 50$	0.7852	0.6317	0.9804	0.0310
genderMale	Gender: Male vs Female	1.0052	0.8577	1.1778	0.9484
raceCaucasian	Race: Caucasian vs African American	0.9644	0.7911	1.1802	0.7226
raceOther	Race: Other vs African American	0.9632	0.5787	1.5459	0.8806
num_medications	Number of medications	1.0127	1.0012	1.0241	0.0294

Term		OR	2.5 %	97.5 %	p.value
time_in_hospital	Length of stay (days)	1.0116	0.9822	1.0414	0.4399
number_emergency	Number of prior emergency visits	1.0731	1.0077	1.1410	0.0254
number_inpatient	Number of prior inpatient visits	1.1769	1.0969	1.2626	0.0000
admission_type_id2	Admission type: Urgent vs Emergency	1.0650	0.8596	1.3140	0.5606
admission_type_id3	Admission type: Elective vs Emergency	0.9027	0.7206	1.1248	0.3670
admission_type_id5	Admission type: Other vs Emergency	0.8498	0.6302	1.1312	0.2749
encounter_index_trunc	Truncated encounter index	0.8054	0.7348	0.8812	0.0000

In the full GLM model, several predictors were significantly associated with 30-day readmission, including age 50–80, number of medications, prior emergency visits, prior inpatient visits, and the truncated encounter index. Based on the odds ratio table, patients aged 50–80 had about 20% lower odds of 30-day readmission compared with those ≤50 years (OR = 0.79, 95% CI: 0.61–0.98). This means that, holding all other variables constant, being in the 50–80 age group is associated with a meaningful reduction in readmission risk. Each additional medication slightly raised the odds of readmission (OR = 1.01, 95% CI: 1.00–1.02), and greater numbers of emergency visits (OR = 1.07, 95% CI: 1.01–1.13) and inpatient visits (OR = 1.18, 95% CI: 1.10–1.27) were associated with higher odds. In contrast, the truncated encounter index showed a strong protective effect; each one-unit increase (up to the cap of 5) corresponded to roughly a 20% reduction in the odds of readmission (OR = 0.81, 95% CI: 0.73–0.88). Confidence intervals for the significant predictors did not cross 1, supporting their statistical significance. InsulinSteady has a p-value slightly above 0.05, suggesting a possible increase in readmission risk, although its 95% confidence interval crossed 1 and therefore the effect was not statistically significant.

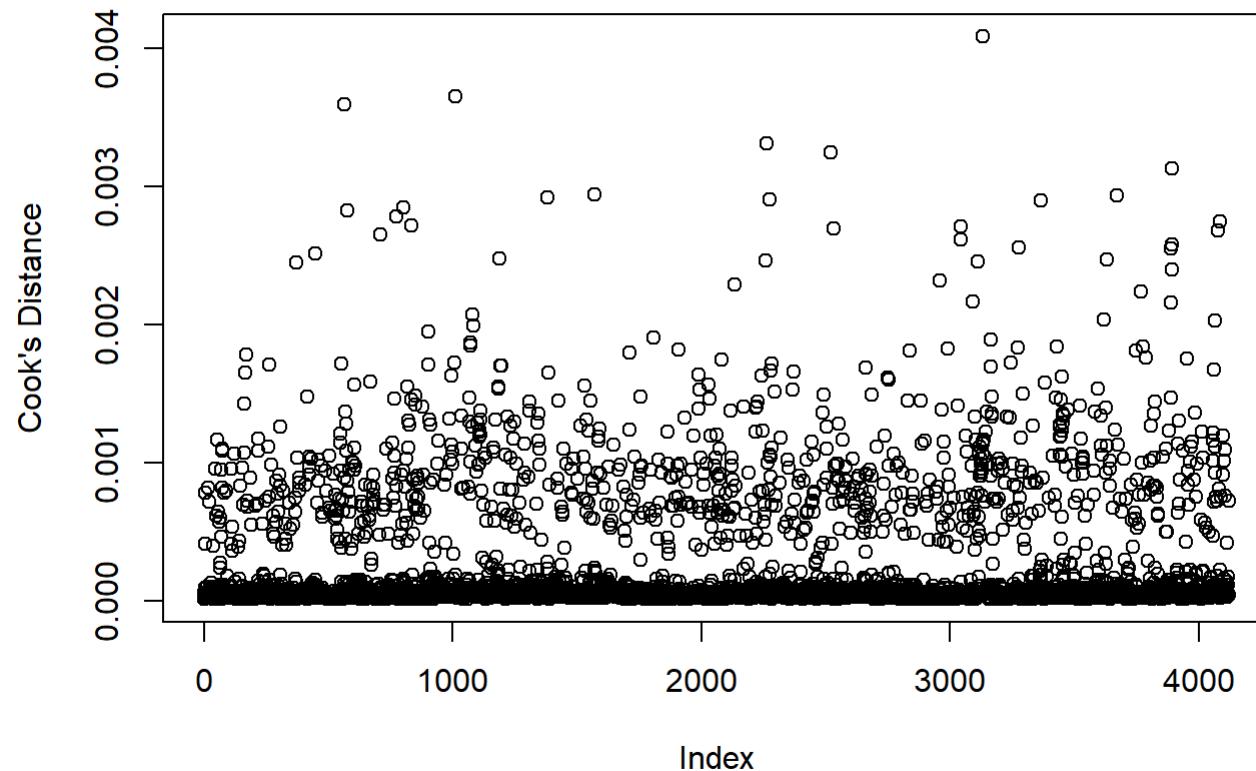
#### Variance Inflation Factors for Logistic GLM

	GVIF	Df	GVIF^(1/(2×Df))
insulin	1.113	3	1.018
age	1.187	2	1.044
gender	1.028	1	1.014
race	1.108	2	1.026
num_medications	1.398	1	1.182

	<b>GVIF</b>	<b>Df</b>	<b>GVIF<sup>(1/(2×Df))</sup></b>
time_in_hospital	1.285	1	1.134
number_emergency	1.226	1	1.107
number_inpatient	2.113	1	1.454
admission_type_id	1.094	3	1.015
encounter_index_trunc	1.983	1	1.408

Based on this variance inflation factor table, all variance inflation factors were low ( $\text{GVIF}^{(1/(2 \times \text{Df}))} \approx 1-2$ ), indicating no evidence of harmful multicollinearity among the predictors. Even the highest values—for the truncated encounter index (1.41) and the number of prior inpatient visits (1.45)—remain well below common concern thresholds (typically 2–5), suggesting that multicollinearity is minimal and unlikely to meaningfully bias the coefficient estimates.

## Cook's Distance for Influential Observations

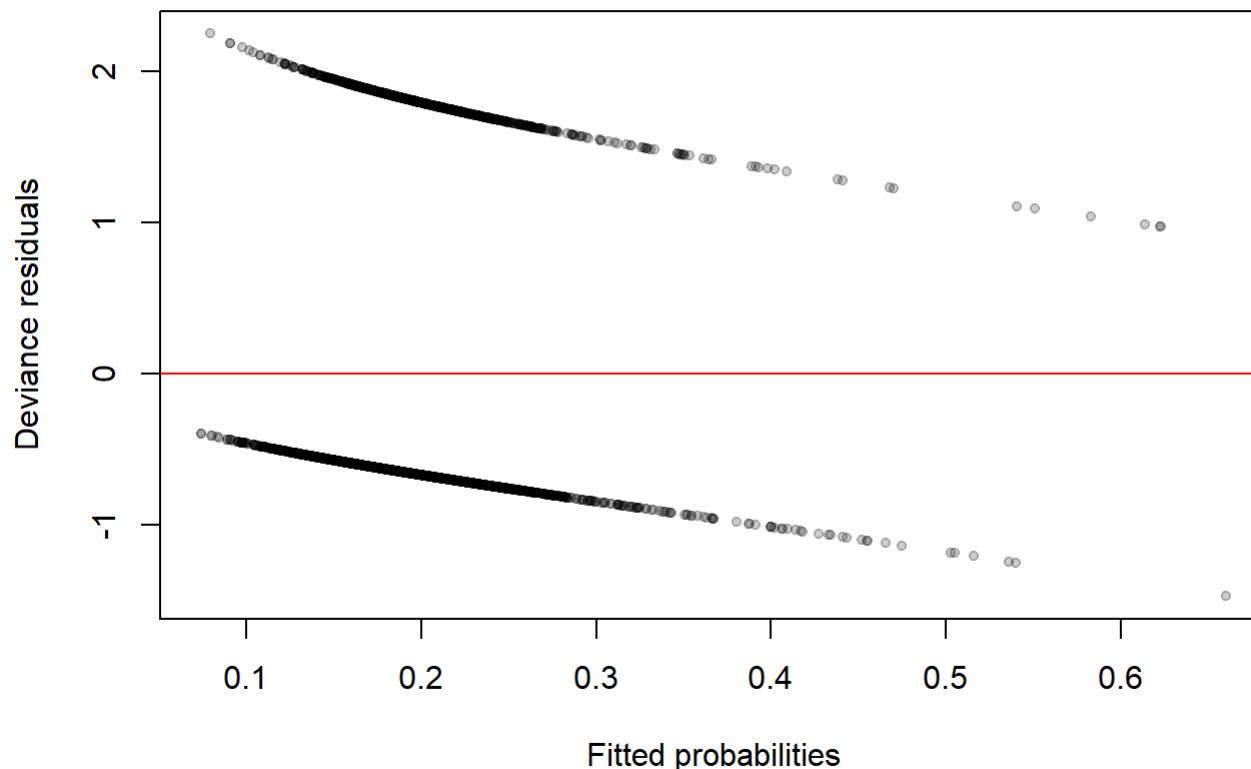


Cook's Distance for Influential Observations

The Cook's distance plot shows that all influence values are extremely small and far below commonly used thresholds. This indicates that no single observation had a notable impact on the fitted GLM.

The dispersion estimate was  $\varphi = 0.97$ , which is very close to 1, indicating no evidence of overdispersion in the logistic GLM.

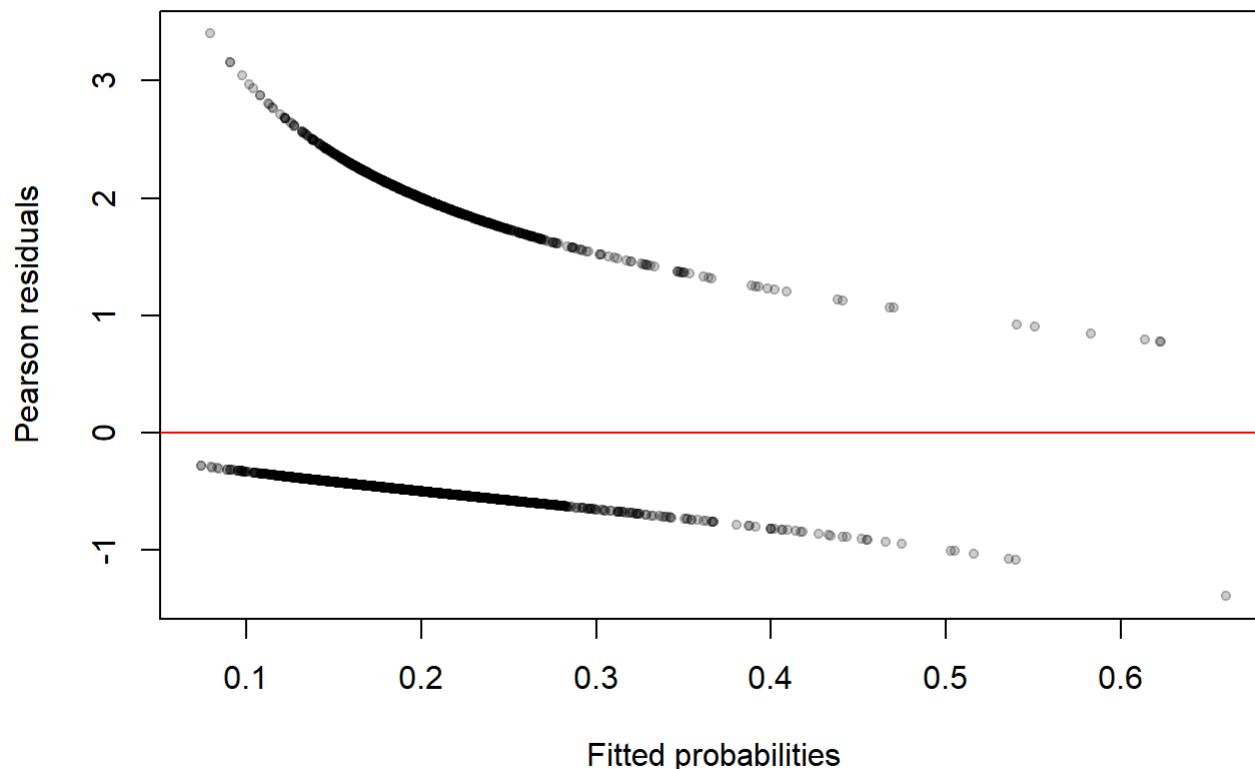
## Deviance Residuals vs Fitted Probabilities



Deviance Residuals vs Fitted Probabilities

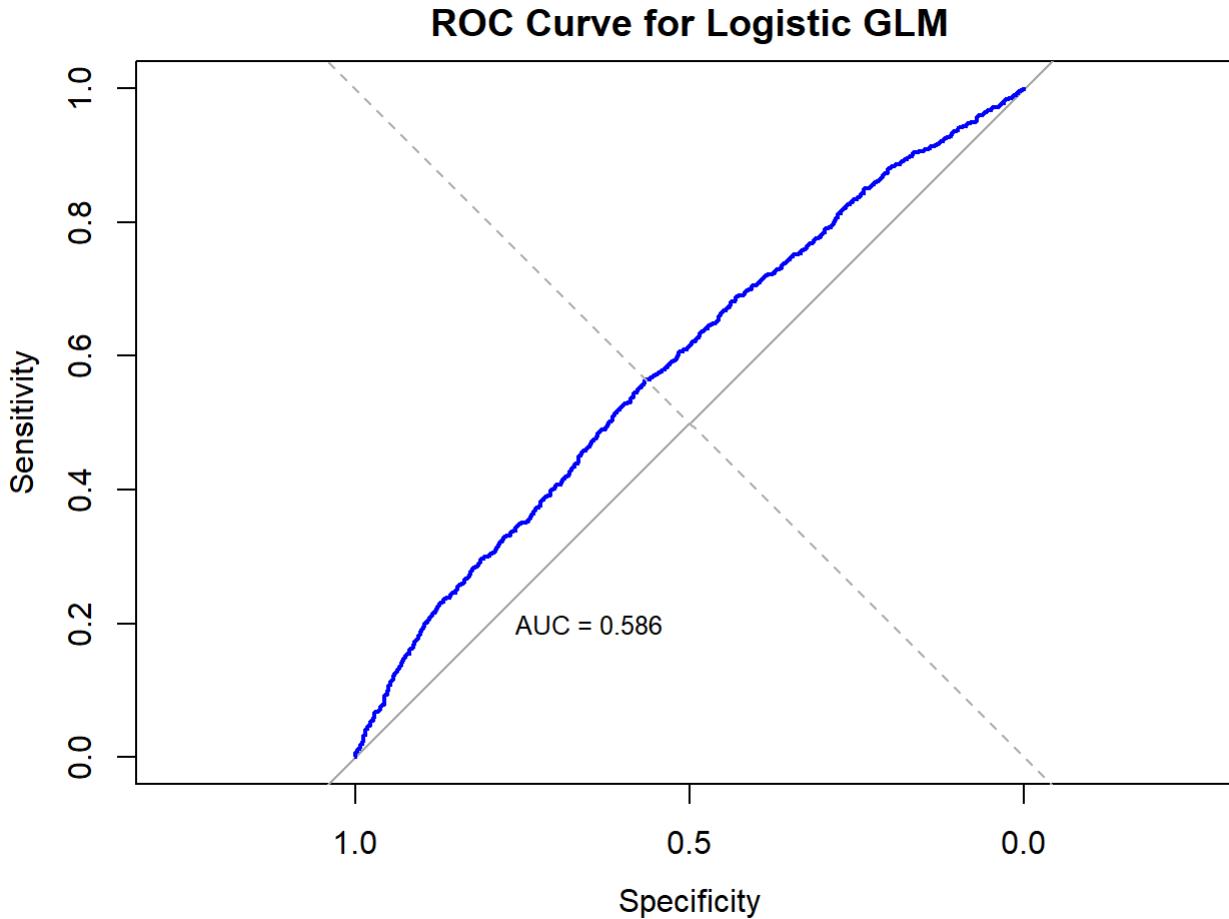
The residual plot of deviance residuals against fitted values shows a symmetric, funnel-shaped pattern typical of logistic regression and no systematic structure, indicating that the model fits reasonably well.

## Pearson Residuals vs Fitted Probabilities



Pearson Residuals vs Fitted Probabilities

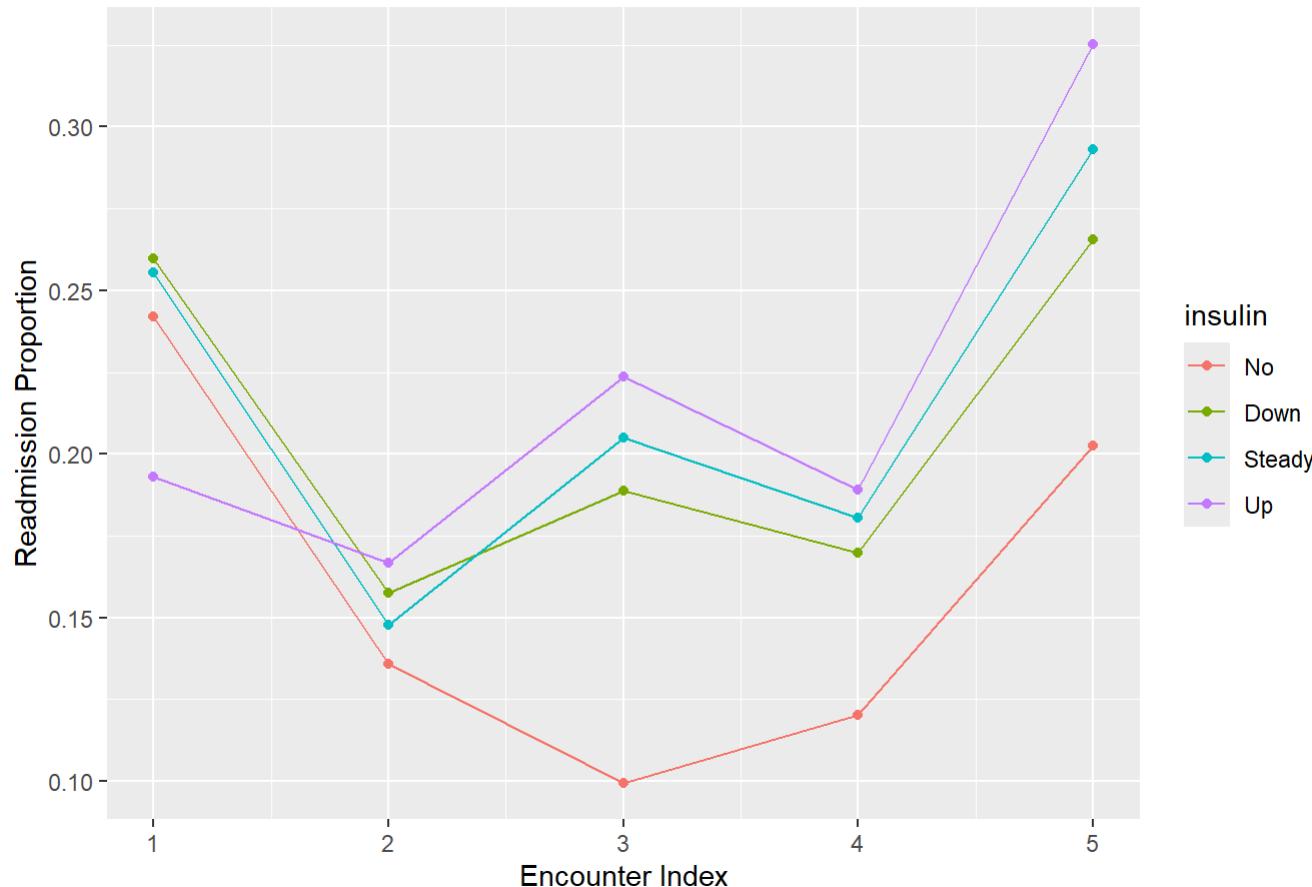
The Pearson residuals against fitted values show the expected curved pattern for logistic regression and no systematic structure, indicating acceptable model fit.



ROC Curve for Logistic GLM

The Hosmer–Lemeshow test gave a p-value of 0.66, indicates no evidence of poor calibration. The predicted and observed readmission rates are reasonably aligned. However, the AUC was 0.586, so the model's ability to distinguish between readmitted and non-readmitted patients is only slightly better than chance.

## Trajectory of Readmission Rates Across Encounters, Stratified by Insulin Regime

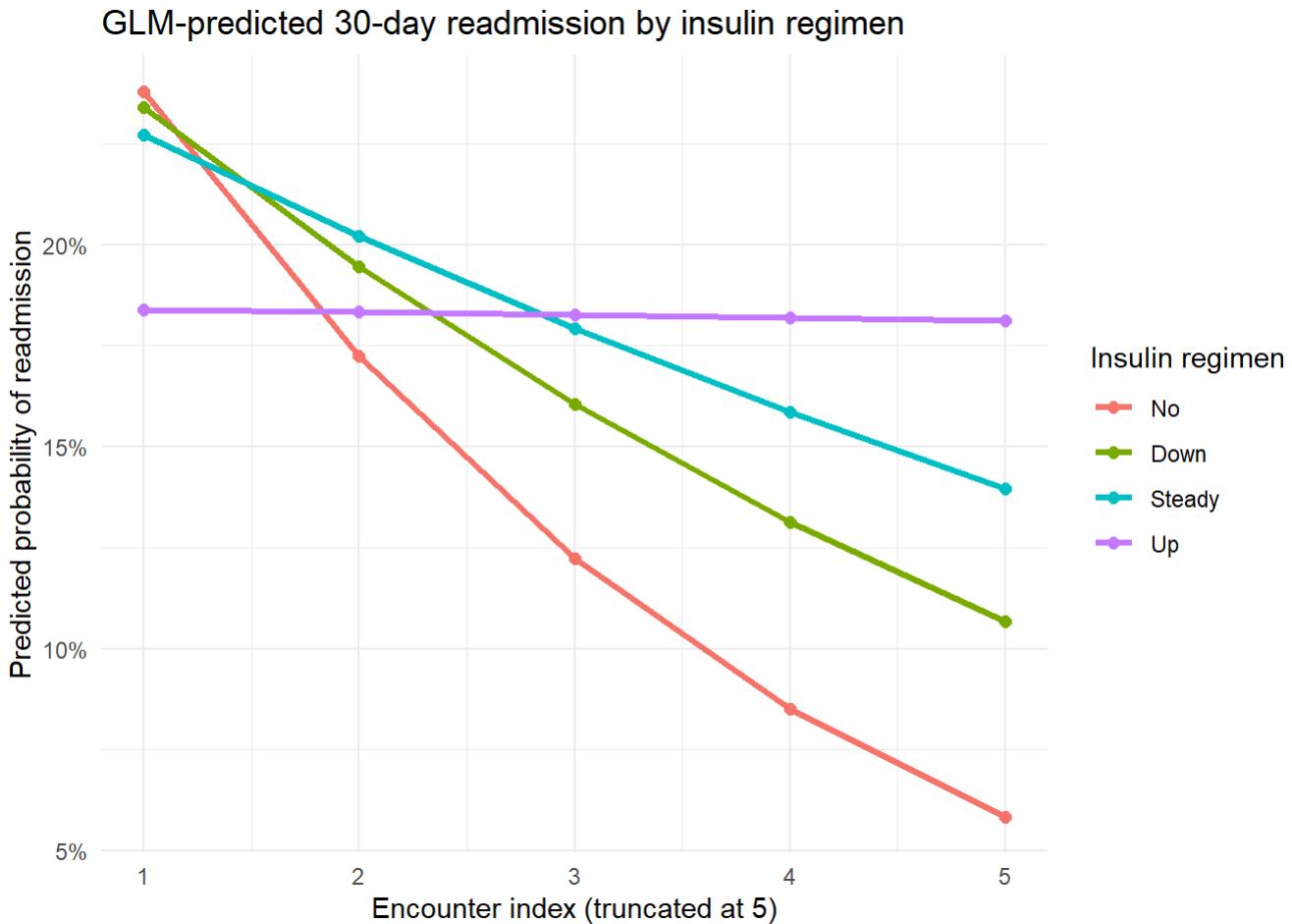


## Trajectory of Readmission Rates Across Encounters, Stratified by Insulin Regimen

The trajectory plot of observed readmission proportions across encounter index and insulin regimen shows clearly non-parallel paths, suggesting that the encounter effect is not the same for all groups.

Including the  $\text{insulin} \times \text{encounter\_index\_trunc}$  interaction improved model fit (AIC decreased from 4026 to 4014), and several interaction terms were statistically significant.  $\text{InsulinSteady} \times \text{Encounter\_index}$  ( $p = 0.003$ ) and  $\text{InsulinUp} \times \text{Encounter\_index}$  ( $p < 0.0001$ ) showed strong evidence that the effect of encounter index differs by insulin group. This indicates that the association between encounter index and readmission is not constant across insulin categories.

The trajectory plot result is consistent with the significant interaction terms in the GLM. Together, they suggest that the slope linking encounter index to readmission probability varies by insulin status. This provides justification for modeling strategies that allow group-specific slopes or correlated repeated measures, rather than relying on a single common slope for all insulin groups.



Model-based predicted 30-day readmission probabilities by encounter index and insulin regimen (GLM with insulin × encounter index interaction)

The model-based GLM prediction curves recapitulate the non-parallel trajectories observed in the raw data, with the "No insulin" group showing a pronounced decline in readmission probability over encounters and the insulin-treated groups following flatter, persistently higher-risk curves. This pattern closely aligns with the GLMM-based

predictions, indicating that the key insulin-by-time interaction is robust to the choice between marginal and mixed-effects modeling frameworks.



Lag-1 Pearson Residual Plot

The lag-1 Pearson residual plot shows clear clustering and diagonal patterns, indicating positive autocorrelation between residuals from consecutive encounters. This violates the independence assumption of the GLM and suggests that within-patient correlation is present, making methods such as GEE or GLMM more appropriate.

Term	Estimate	SE_Model	p_Model	SE_Robust	p_Robust
(Intercept)	-1.3158	0.1731	0.0000	0.1860	0.0000
insulinDown	0.1194	0.1245	0.3379	0.1291	0.3554
insulinSteady	0.1828	0.0958	0.0564	0.0961	0.0572
insulinUp	0.1489	0.1266	0.2396	0.1258	0.2365
age>=80	-0.1354	0.1362	0.3201	0.1557	0.3845
age50-80	-0.2418	0.1121	0.0310	0.1316	0.0662
genderMale	0.0052	0.0809	0.9484	0.0867	0.9518
raceCaucasian	-0.0362	0.1020	0.7226	0.1087	0.7390
raceOther	-0.0375	0.2496	0.8806	0.2225	0.8662
num_medications	0.0126	0.0058	0.0294	0.0061	0.0401
time_in_hospital	0.0115	0.0149	0.4399	0.0144	0.4239
number_emergency	0.0706	0.0316	0.0254	0.0344	0.0403
number_inpatient	0.1629	0.0359	0.0000	0.0341	0.0000
admission_type_id2	0.0630	0.1082	0.5606	0.1071	0.5565
admission_type_id3	-0.1024	0.1135	0.3670	0.1104	0.3538
admission_type_id5	-0.1627	0.1490	0.2749	0.1417	0.2508
encounter_index_trunc	-0.2164	0.0463	0.0000	0.0512	0.0000

Cluster-robust standard errors were computed to adjust for within-patient correlation across repeated encounters. While coefficient estimates were unchanged, several standard errors increased. Notably, the age 50–80 category was no longer statistically significant under the robust SEs ( $p = 0.066$ , from the table comparing model-based and cluster-robust standard errors), indicating that the naïve GLM slightly underestimated uncertainty. The main conclusions for number of medications, emergency visits, inpatient visits, and the encounter index remained robust.

## GEE

Generalized Estimating Equations (GEE) provide an alternative framework for analyzing correlated binary outcomes that targets population-averaged effects while accounting for within-cluster correlation through a working correlation structure. In the context of our data, a GEE approach would treat repeated encounters within the same patient as a cluster and estimate marginal log-odds ratios for insulin regimen, encounter history, and their interaction, with robust ("sandwich") standard errors that remain valid even if the working correlation is misspecified.

Conceptually, the GEE model occupies a middle ground between the independence GLM with cluster-robust standard errors and the subject-specific GLMM. Like the robust GLM, GEE emphasizes average effects at the population level rather than individual patient trajectories. At the same time, it explicitly encodes correlation through, for example, an exchangeable or autoregressive working structure rather than relying solely on cluster-robust variance estimates. Because our primary scientific questions focused on patient-level heterogeneity in baseline risk and rate of change over time, as well as on decomposing variance into within- and between-patient components, we elected to emphasize the GLMM as the main longitudinal modeling strategy.

To connect the subject-specific GLMM estimates back to a population-averaged interpretation analogous to GEE, we applied Zeger's attenuation method to approximate marginal effects from the random-slope model. The close agreement between conditional and approximate marginal odds ratios suggests that, in this data set, the choice between a GEE-type marginal model and a mixed-effects model is unlikely to materially alter substantive conclusions. We therefore view the GLMM and the marginal GLM with robust standard errors as complementary perspectives on the same underlying associations, with the former providing richer insight into individual trajectories.

## GLMM

### Exploratory plots

Four numeric covariates were standardized to facilitate model convergence and improve the numerical stability of the optimization algorithm. Use centered time (`encounter_index_trunc_c = encounter_index_trunc - mean(encounter_index_trunc)`) for easier interpretation of main effects.

### Fit random intercept model

We employed adaptive Gauss-Hermite quadrature with 10 integration points ( $nAGQ = 10$ ) rather than the standard Laplace approximation ( $nAGQ = 1$ ). For binary response data, the standard Laplace approximation is known to potentially underestimate the variance of random effects and introduce bias into the fixed effects coefficients. Increasing the number of quadrature points provides a more accurate approximation of the log-likelihood, thereby yielding more reliable parameter estimates.

We need LRT to check the interaction term

To assess whether the effect of time (encounter index) on readmission risk varies by insulin status, we compared the full model (including the interaction term) against a reduced model (main effects only) using a Likelihood Ratio Test (LRT).

The LRT indicated that the inclusion of the interaction term `insulin * encounter_index_trunc_c` resulted in a statistically significant improvement in model fit ( $\chi^2(3) = 16.11, p < 0.005$ ). Furthermore, the Akaike Information Criterion (AIC) decreased substantially from 4019.0 in the reduced model to 4008.9 in the full model, confirming that the interaction model provides a superior fit to the data. This suggests that the trajectory of readmission risk over time is significantly different depending on the patient's insulin regimen.

## Random Slopes Model

$nAGQ = 10$  can't be used in the random slope model, so we use  $nAGQ = 1$  here.

We evaluated the necessity of including random slopes for the time effect (`encounter_index_trunc_c`) to allow for inter-individual variation in the trajectory of readmission risk. A Likelihood Ratio Test comparing the random-slope model (`m3`) against the random-intercept-only model (`m1`) showed a statistically significant improvement in model fit ( $\chi^2(2) = 14.42, p < 0.001$ ). Additionally, the random-slope model exhibited a lower AIC (3998.5 vs. 4008.9), indicating a better trade-off between goodness of fit and model complexity. Consequently, we selected the random-slope model (`m3`) as the final model. This suggests that not only do baseline risks vary across patients, but the rate of change in readmission risk over time also varies significantly between individuals.

Model comparison for mixed-effects specifications of 30-day readmission

Model	Specification	AIC	Decision	LRT $\chi^2$ (vs simpler)	p-value

Model	Specification	AIC	Decision	LRT $\chi^2$ (vs simpler)	p-value
M0: RI, main effects	Random intercept; no Insulin $\times$ Time	4019.033	Reference	-	-
M1: RI, + Insulin $\times$ Time	Random intercept; Insulin $\times$ Time	4008.928	Keep interaction	16.1	0.00108
M2: RS, + Insulin $\times$ Time	Random intercept + random slope; Insulin $\times$ Time	3998.512	Final model	14.42	0.000741

## Interaction Effects

There was a significant interaction between insulin status and the timing of encounters. For the reference group (patients not on insulin), the probability of 30-day readmission showed a steep and significant decline over successive visits ( $\beta_{time} = -0.462, p < 0.001$ ). However, this protective temporal trend was significantly attenuated for patients on steady insulin ( $\beta_{interaction} = 0.232, p = 0.014$ ) and further diminished for patients whose insulin dosage was increased ( $\beta_{interaction} = 0.329, p = 0.004$ ). This indicates that while Insulin Up patients did experience a marginal reduction in risk over time, their **rate of recovery was significantly slower and flatter** compared to patients not requiring insulin therapy.

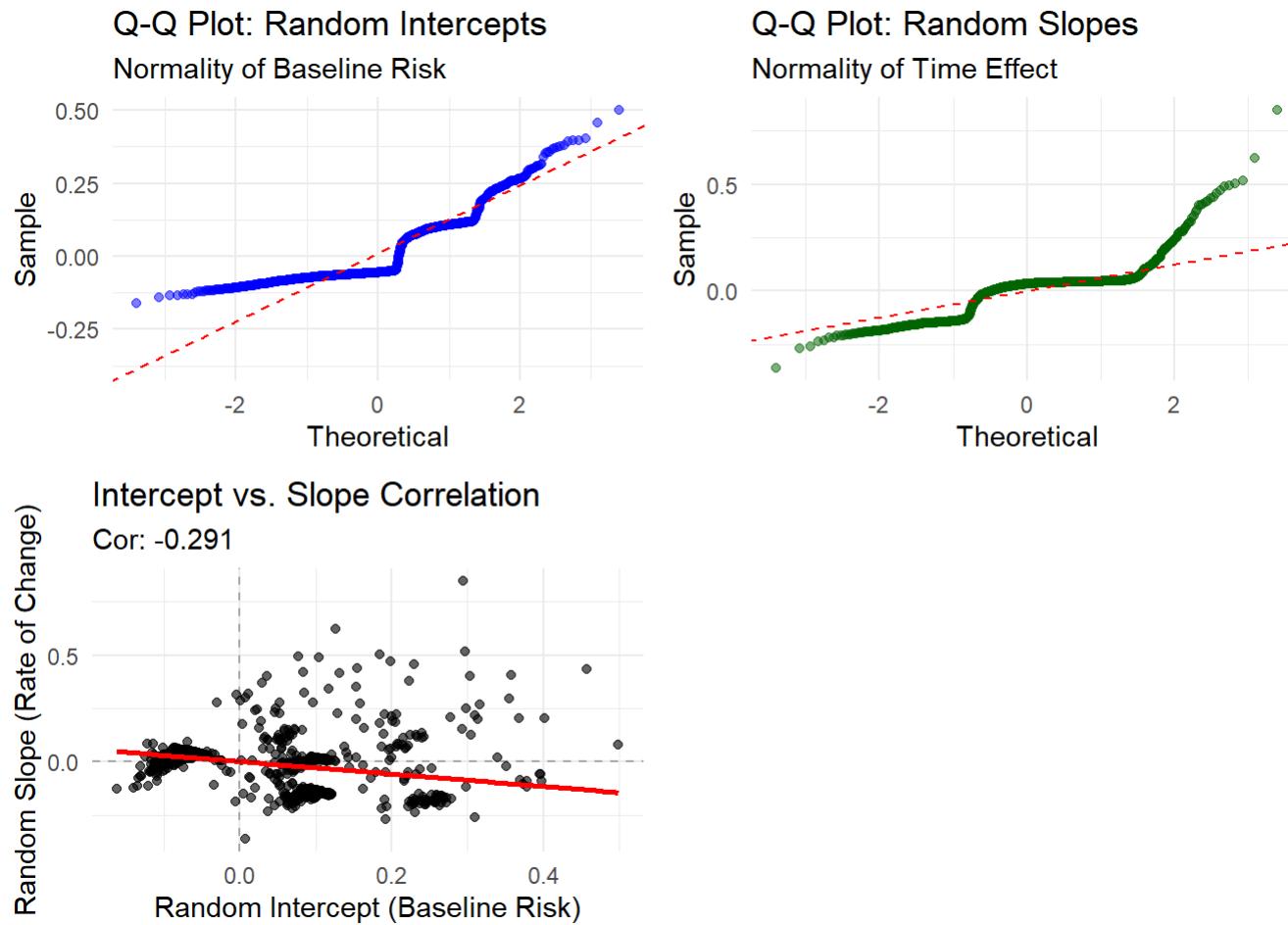
## ICC and Variance

The Adjusted Intraclass Correlation Coefficient (ICC) was 0.113, indicating that approximately 11.3% of the latent variation in readmission risk is attributable to unobserved heterogeneity between patients (such as unique genetics, chronic frailty, or compliance behavior). The remaining variation is driven by within-patient time-varying factors or random noise. An ICC exceeding 10% confirms substantial clustering within subjects, validating the necessity of the mixed-effects framework to account for non-independence in the data.

## Covariance Matrix

The model's direct Maximum Likelihood Estimation (MLE) of the covariance matrix yielded a correlation coefficient of only  $\rho = -0.04$ . This parameter represents the unbiased population-level relationship. A correlation of -0.04 is statistically negligible, indicating that in the true underlying population, a patient's baseline risk level is effectively independent of their rate of change over time.

## Diagnostic



Random intercept and slope diagnostics for GLMM m3

**Evaluation of Random Effects Structure:** The diagnostic plots for the random effects structure (Model m3) revealed deviations from ideal normality for both random intercepts and random slopes. The Q-Q plots exhibit a "stepped" pattern, which is a characteristic artifact in generalized linear mixed models (GLMMs) when the outcome is binary and cluster sizes are small (i.e., limited visits per patient).

Despite these deviations, two key insights emerged: 1. Correlation of Effects: We observed a moderate negative correlation ( $\rho = -0.291$ ) between the random intercepts and slopes. This suggests a regression to the mean

dynamic: patients with higher baseline risks (higher intercepts) tend to experience a sharper decline in risk over time (more negative slopes), whereas lower-risk patients exhibit more stable trajectories. 2. Model Robustness: Although the random effects distributions are not perfectly normal, previous simulation studies in GLMMs have demonstrated that fixed effect estimates remain robust to misspecification of the random effects distribution, especially given our large sample size. Therefore, retaining the random-slope structure provides a more realistic representation of patient heterogeneity without compromising the validity of the fixed effect inferences.

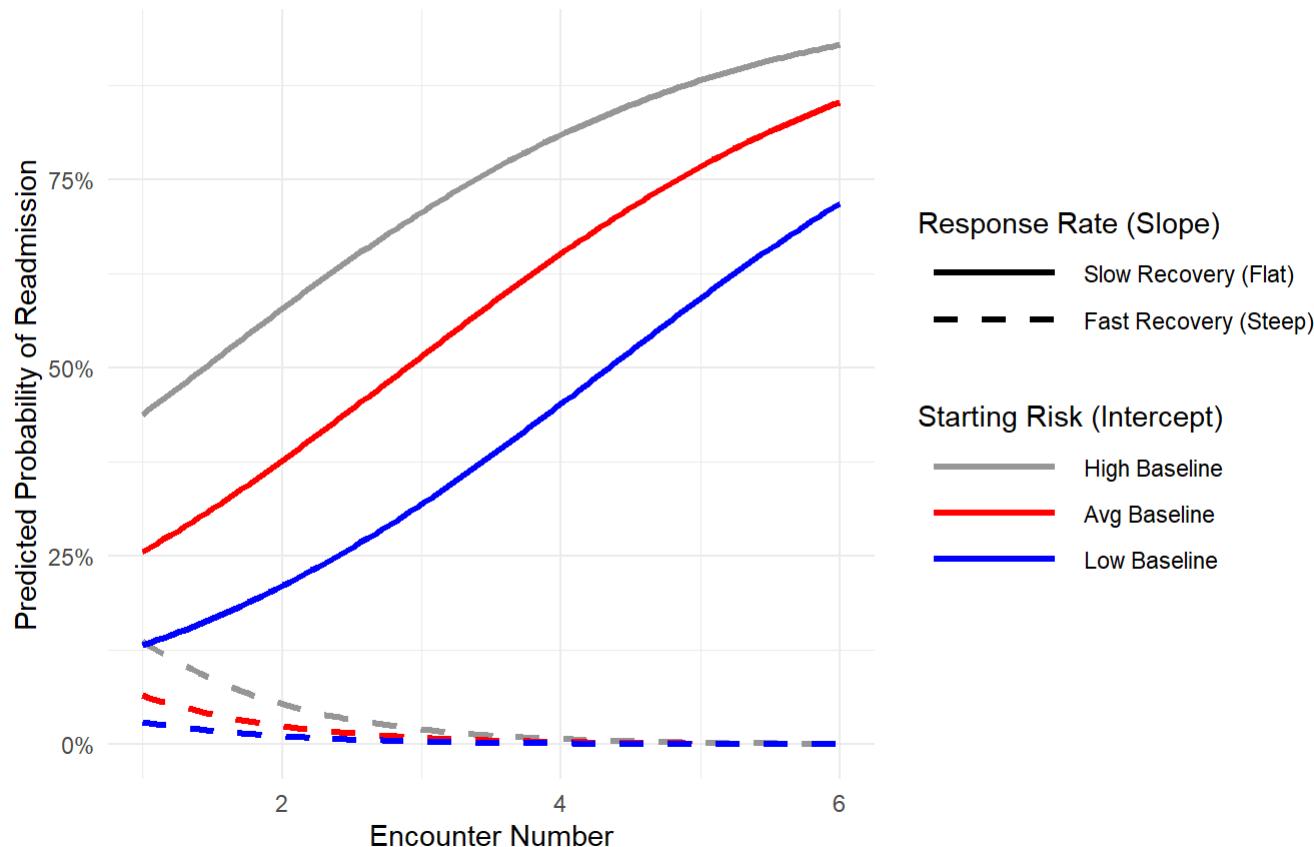
The discrepancy between the visual estimate (-0.291) and the model estimate (-0.04) is explained by the '**Shrinkage Effect**' common in GLMMs. Because many patients have few visits, the model 'shrinks' their uncertain individual estimates toward the population mean, which mathematically forces a spurious negative correlation in the scatterplot. **Therefore, we rely on the model output (-0.04) as the valid inference.** Our final clinical conclusion is that a patient's starting frailty (intercept) does not predict how fast they will stabilize (slope); these are two distinct biological characteristics.

We assessed multicollinearity using the Generalized Variance Inflation Factor (GVIF). All predictors showed low GVIF values, falling well below the conservative threshold of 5.

## Simulation

## Patient Heterogeneity: Baseline vs. Recovery Speed

Decoupling 'Where they start' from 'How fast they improve' (Model m3)

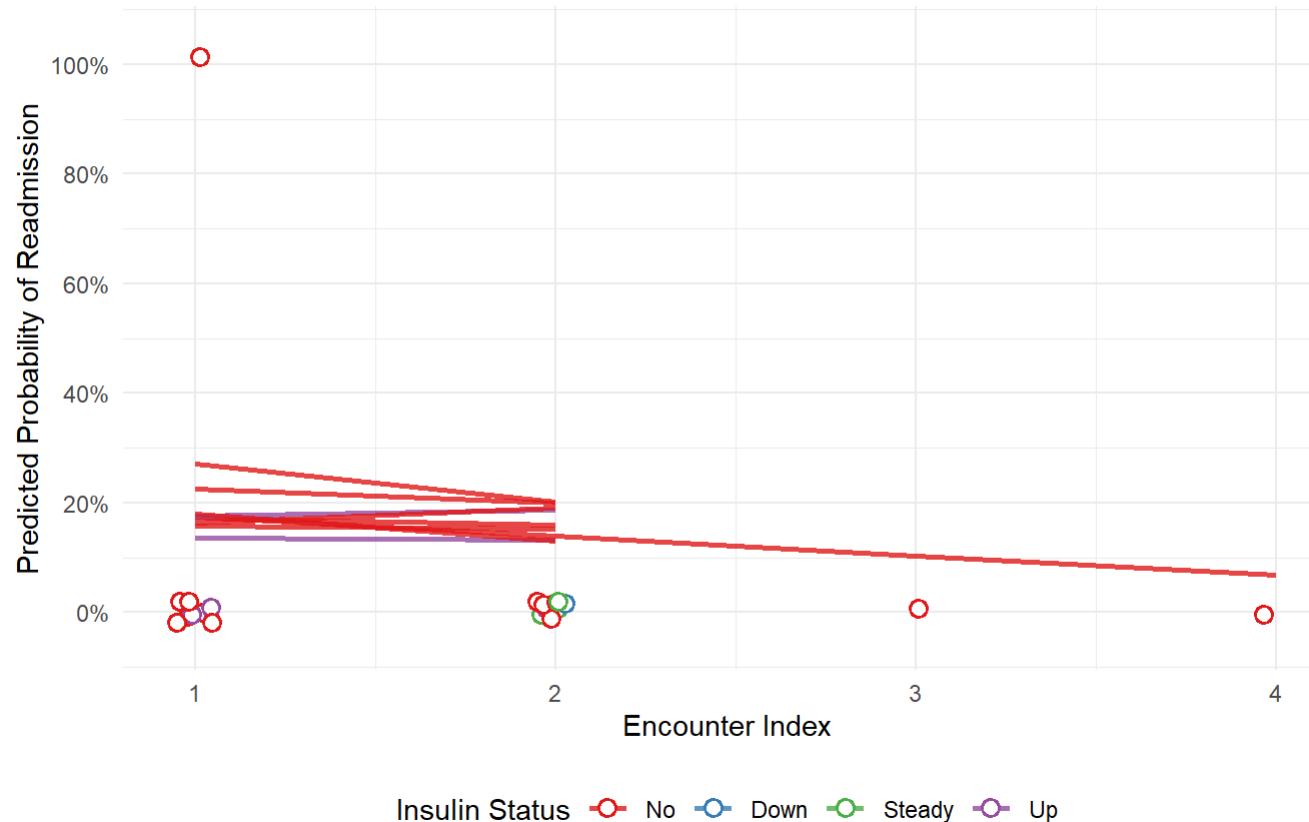


Simulated readmission trajectories under varying baseline risk and recovery speed (Random-slope GLMM m3)

Our Random Slope model (m3) reveals a more nuanced reality because the correlation between baseline and slope is near zero. A simulation plot illustrates six illustrative scenarios defined by two clinically meaningful dimensions. The color encodes starting risk (high-, average-, and low-risk baselines), and the line type encodes recovery speed (fast versus slow decline in risk over encounters). Among otherwise similar high-risk patients, those with fast recovery exhibit a steep drop in readmission probability over visits, whereas those with slow recovery remain persistently high-risk despite repeated encounters. Distinguishing between these patterns is critical for targeting limited care-management resources toward patients whose trajectories are resistant to improvement.

## Subject-Specific Predicted Trajectories (Sample n=9)

Lines = GLMM fitted values (BLUPs); Points = Observed Outcomes (0=No, 1=Yes)



Subject-specific predicted trajectories for a random sample of nine patients (Random-slope GLMM m3)

Conditional (GLMM) vs. Approximate Marginal (Zeger) Effects

term	Cond_Beta	Cond_OR	Marg_Beta_Approx	Marg_OR_Approx
insulinDown	0.170	1.186	0.165	1.180
insulinSteady	0.239	1.270	0.232	1.261
insulinUp	0.121	1.129	0.118	1.125
encounter_index_trunc_c	-0.462	0.630	-0.449	0.638

term	Cond_Beta	Cond_OR	Marg_Beta_Approx	Marg_OR_Approx
time_in_hospital	0.046	1.047	0.044	1.045
num_medications	0.096	1.100	0.093	1.097
number_emergency	0.083	1.087	0.081	1.084
number_inpatient	0.152	1.164	0.148	1.159
admission_type_id2	0.063	1.065	0.061	1.063
admission_type_id3	-0.107	0.899	-0.104	0.901
admission_type_id5	-0.147	0.864	-0.142	0.867
age<=50	0.212	1.236	0.206	1.229
age>=80	0.130	1.139	0.126	1.134
genderMale	-0.012	0.988	-0.012	0.989
raceCaucasian	-0.037	0.964	-0.036	0.965
raceOther	0.004	1.004	0.004	1.004
insulinDown:encounter_index_trunc_c	0.126	1.134	0.122	1.130
insulinSteady:encounter_index_trunc_c	0.232	1.261	0.225	1.253
insulinUp:encounter_index_trunc_c	0.329	1.389	0.319	1.376

## Results: Analysis of 30-Day Readmission Risk (Random-Slope GLMM)

### 1. Model Specification and Estimation

To formally account for correlation among repeated encounters within patients, we fitted a logistic Generalized Linear Mixed Model (GLMM). Let  $Y_{ij}$  denote the 30-day readmission indicator for patient  $i$  at encounter  $j$ , and let  $\text{Time}_{ij}$  denote the centered, truncated encounter index. The conditional log-odds of readmission are modeled as

$$\text{logit}\{\Pr(Y_{ij} = 1 | \mathbf{b}_i)\} = \beta_0 + \boldsymbol{\beta}_{\text{Ins}}^\top \mathbf{Z}_{\text{Ins},ij} + \beta_{\text{Time}} \text{Time}_{ij} + \boldsymbol{\beta}_{\text{Int}}^\top (\mathbf{Z}_{\text{Ins},ij} \times \text{Time}_{ij}) + \boldsymbol{\gamma}^\top \mathbf{X}_{ij} + b_{0i} + b_{1i} \text{Time}_{ij},$$

where  $\mathbf{Z}_{\text{Ins},ij}$  encodes insulin regimen,  $\mathbf{X}_{ij}$  collects additional covariates (length of stay, number of medications, prior emergency and inpatient visits, admission type, age group, gender, and race), and  $\mathbf{b}_i = (b_{0i}, b_{1i})^\top$  are

patient-specific random effects. We assumed a bivariate normal distribution for the random effects,

$$\mathbf{b}_i \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix}\right),$$

so that  $\sigma_0^2$  and  $\sigma_1^2$  represent between-patient variability in baseline risk and time trends, respectively, and  $\rho$  captures the correlation between baseline risk and slope. Under this parameterization, an approximate marginal intraclass correlation coefficient (ICC) for the latent propensity to be readmitted can be defined as

$$\text{ICC} = \frac{\sigma_0^2}{\sigma_0^2 + \pi^2/3},$$

where  $\pi^2/3$  is the variance of the standard logistic distribution. Fixed effects were estimated by maximum likelihood using adaptive Gauss–Hermite quadrature for the random-intercept model and the Laplace approximation for the random-slope model.

Model selection proceeded in stages. We first compared a random-intercept model with and without the insulin-by-time interaction using a Likelihood Ratio Test (LRT) and Akaike Information Criterion (AIC), then compared the best-fitting random-intercept model against a random-slope specification. The inclusion of the insulin-by-time interaction significantly improved model fit ( $\chi^2(3) = 16.11, p < 0.005$ , AIC: 4008.9 vs. 4019.0). Allowing random slopes for time further improved fit relative to the random-intercept model ( $\chi^2(2) = 14.42, p < 0.001$ , AIC: 3998.5 vs. 4008.9), indicating substantial between-patient heterogeneity in both baseline risk and trajectories over time. Consequently, the random-slope GLMM (model [m3](#)) was adopted as the primary longitudinal model.

## 2. Random Effects and Subject Heterogeneity

The final model revealed substantial patient heterogeneity. The variance of the random intercept was **0.175** (SD = 0.418), indicating significant differences in initial frailty across patients. The variance of the random slope was **0.165** (SD = 0.406), confirming that patients recover (or deteriorate) at significantly different rates. The correlation between random intercepts and slopes was negligible ( $\rho = -0.04$ ), suggesting that a patient's baseline risk is largely independent of their rate of recovery. The Adjusted Intraclass Correlation Coefficient (ICC) was **0.113**, implying that approximately **11.3%** of the variation in readmission risk is attributable to unobserved subject-specific characteristics.

## 3. Fixed Effects: The Interaction of Insulin and Time

The analysis demonstrated a significant interaction between insulin status and the timing of encounters. Protective Trend in Reference Group: For patients not on insulin, the probability of readmission decreased steeply with successive visits (Cond. OR = 0.630,  $\beta = -0.462$ ,  $p < 0.001$ ), indicating a strong 'learning' or stabilization effect. Attenuation in Insulin Groups: This protective temporal trend was significantly blunted for patients on insulin. The interaction term for the Insulin Steady group was significant ( $\beta_{interaction} = 0.232$ ,  $p = 0.014$ ), and even more pronounced for the Insulin Up group ( $\beta_{interaction} = 0.329$ ,  $p = 0.004$ ). While 'Insulin Up' patients still experienced a slight decrease in risk over time (Net slope  $\approx -0.13$ ), their rate of improvement was significantly slower compared to the reference group.

#### 4. Conditional vs. Approximate Marginal Effects (Zeger Method)

Since GLMM parameters represent subject-specific (conditional) effects, we applied Zeger's method to approximate population-averaged (marginal) effects. The Zeger attenuation factor was **1.03**, resulting in a modest **2.89% reduction** in effect sizes when scaling from the individual to the population level. The Conditional ORs and Marginal ORs remained highly consistent (e.g., Insulin Up Cond. OR = 1.129 vs. Marg. OR = 1.125). This suggests that despite patient heterogeneity, the observed risk factors apply robustly to the population average.

#### 5. Other Clinical Predictors

Controlling for the time-insulin interaction, other clinical factors showed the following associations: Prior inpatient visits remained a significant predictor, with each additional visit increasing the odds of readmission by 16.4% (Cond. OR = 1.164,  $p = 0.019$ ). The number of medications showed a borderline significant association (Cond. OR = 1.100,  $p = 0.051$ ), serving as a likely proxy for comorbidity complexity. Age, gender, and race were not statistically significant in this adjusted longitudinal framework.

---

### Summary Conclusion

In conclusion, the **Random Slope model** provides superior fit, revealing that patients differ not only in their baseline risk but also in their recovery trajectories. While the "No Insulin" group exhibits a rapid decline in readmission risk over repeated encounters, patients requiring insulin dose escalation ("Insulin Up") show a significantly **flatter recovery curve**. This suggests that insulin intensification serves as a marker for a more resistant, high-risk phenotype that benefits less from the standard "stabilization over time" effect.

### Comparison

We compared a series of modeling strategies that make different assumptions about within-patient correlation and target different inferential estimands. The baseline analysis used a standard logistic GLM treating encounters as independent. This model provided interpretable population-averaged odds ratios and adequate global fit (Hosmer–Lemeshow  $p = 0.66$ ), but the modest discrimination ( $AUC \approx 0.59$ ) and evidence of serial correlation in the lag-1 Pearson residual plot indicated that the independence assumption was violated. Cluster-robust (sandwich) standard errors partially addressed this issue by inflating uncertainty for clustered observations without altering point estimates; several covariates, such as age 50–80, lost nominal statistical significance under robust SEs, highlighting that the naïve GLM underestimated uncertainty.

We then fitted a subject-specific Generalized Linear Mixed Model (GLMM) with random intercepts and, subsequently, random slopes for encounter index. Compared with the marginal GLM, the GLMM yielded effect estimates of similar magnitude for key clinical predictors (e.g., prior inpatient utilization, number of medications), but revealed substantial between-patient heterogeneity in both baseline risk and trajectories over time. The variance components (intercept SD  $\approx 0.42$ , slope SD  $\approx 0.41$ ) and an adjusted ICC of 0.113 showed that approximately 11% of the variability in readmission risk was attributable to unobserved patient-level factors. Likelihood ratio tests and AIC clearly favored the random-slope model over both the random-intercept-only GLMM and the marginal GLM, indicating that allowing patient-specific time trends meaningfully improves model fit.

From an inferential perspective, the GLMM focuses on subject-specific (conditional) effects, while the GLM with cluster-robust standard errors targets population-averaged effects. Using Zeger's attenuation factor ( $\sim 1.03$ ), we showed that conditional and approximate marginal odds ratios from the GLMM were nearly identical, suggesting that the substantive conclusions are robust to the choice of estimand. In particular, the insulin-by-time interaction remained significant and directionally consistent across all frameworks: patients in the "No insulin" group experienced a steep decline in readmission risk with successive encounters, whereas those with insulin dose escalation ("Insulin Up") showed a markedly attenuated improvement. Overall, the comparison indicates that while simpler marginal models capture the main average effects, the random-slope GLMM is better suited to describing clinically relevant heterogeneity in longitudinal readmission trajectories.

Taken together, these results highlight clear trade-offs among the candidate models. The independence GLM is simple, computationally efficient, and yields easily interpretable population-averaged odds ratios, but it rests on an implausible independence assumption and underestimates uncertainty when within-patient correlation is present. The GLM with cluster-robust standard errors relaxes the variance assumptions and provides robust

population-averaged inference, yet it cannot decompose variation into within- and between-patient components or describe individual trajectories. The random-intercept GLMM captures between-patient heterogeneity in baseline risk but still imposes a common time trend for all subjects. By contrast, the random-slope GLMM allows both baseline risk and the rate of change over time to vary by patient, accommodates the observed correlation structure, and, after Zeger attenuation, yields marginal effects that closely align with the robust GLM. For these reasons, we base our primary clinical interpretations and recommendations on the random-slope GLMM, using the marginal GLM with cluster-robust standard errors as a complementary sensitivity analysis for population-averaged effects.

## Discussion

In this longitudinal analysis of diabetic inpatients, we found that 30-day readmission risk declined meaningfully over successive encounters, but the magnitude of this improvement depended strongly on insulin regimen. Among patients not treated with insulin, repeated encounters were associated with a steep reduction in the odds of readmission, consistent with a stabilization or “learning” effect as care became more coordinated over time. In contrast, patients requiring insulin intensification, particularly those in the “Insulin Up” group, exhibited a much flatter trajectory: their readmission risk started higher and declined more slowly, even after adjusting for prior utilization, admission type, and comorbidity burden. Approximately 11% of the total variability in readmission risk was attributable to unobserved patient-level heterogeneity, underscoring the importance of modeling repeated measures at the individual level.

These findings have several clinical implications. First, insulin intensification appears to function less as a simple treatment decision and more as a marker of a high-risk phenotype characterized by metabolic instability and treatment complexity. Such patients not only start at a higher baseline risk of readmission but also benefit less from the “usual” gains that accrue with ongoing contact with the health system. Second, the strong and persistent effects of prior inpatient and emergency visits indicate that recent healthcare utilization captures a dimension of frailty and system-level vulnerability that is not fully explained by demographic characteristics. Taken together, these results suggest that readmission reduction efforts should prioritize patients undergoing insulin escalation and those with heavy recent utilization, for example through more intensive discharge planning, early outpatient follow-up, or structured insulin titration programs.

Methodologically, our study illustrates the value of explicitly accounting for within-patient correlation in longitudinal electronic health record data. Although the marginal GLM with cluster-robust standard errors

produced broadly similar point estimates and adequate calibration, diagnostics revealed serial correlation that violated the independence assumption. The random-slope GLMM not only improved statistical fit but also provided a more clinically meaningful decomposition of variation into baseline risk and trajectory over time. The close agreement between conditional and approximate marginal odds ratios implies that our substantive conclusions are not an artifact of the modeling framework, but rather reflect stable underlying associations between insulin regimen, encounter history, and readmission risk.

Several limitations should be considered when interpreting these results. First, the analysis relies on routinely collected administrative and clinical data, which are subject to coding error and lack detailed information on disease severity, glycemic control (e.g., HbA1c), outpatient adherence, and socioeconomic determinants of health. Residual confounding by these unmeasured factors is likely. Second, patients contributed a limited number of encounters, and the binary outcome led to non-ideal normality of the random effects, as suggested by the stepped Q–Q plots. While prior work indicates that GLMM fixed effects are robust to moderate misspecification of the random-effects distribution, this remains a potential source of bias. Third, the modest AUC (~0.59) highlights that our models, though useful for inference, are not suitable as stand-alone clinical prediction tools; future work should incorporate richer covariate information and possibly flexible machine-learning components if prediction is the primary goal.

Despite these limitations, our study leverages a large, real-world cohort with repeated encounters per patient, employs rigorous diagnostic checks, and compares complementary modeling frameworks for correlated binary outcomes. The convergence of evidence across models strengthens confidence in the central conclusion: while most patients experience declining readmission risk over time, those requiring insulin escalation constitute a persistently vulnerable subgroup that does not fully “catch up” despite repeated contact with the healthcare system. Targeting this subgroup for enhanced transitional care and proactive insulin management may yield disproportionate gains in reducing potentially preventable readmissions. Future research should investigate whether interventions tailored to insulin-intensified patients can modify the trajectory we observed and explore how integrating patient-level behavioral and social data further improves risk stratification.

## Conclusion

In summary, we provide a longitudinal characterization of 30-day readmission risk among patients with diabetes that integrates marginal and subject-specific modeling approaches. Readmission risk generally declines with repeated encounters, but this improvement is substantially attenuated in patients undergoing insulin escalation,

who remain at persistently elevated risk even after accounting for prior utilization and other covariates. By demonstrating that conclusions are robust across GLM with robust standard errors and random-slope GLMMs, while highlighting clinically meaningful heterogeneity in individual trajectories, our work underscores the need for targeted, patient-centered strategies focused on insulin-intensified and high-utilization patients to effectively reduce preventable readmissions.