Change in the Mean Total Score of the Patient Health Questionnaire-9 among Bipolar Depression Patients treated with Lumateperone in the United States: An Electronic Health Records Study

Madhav Namjoshi[1], Danae A. Black[2], Lawrence Rasouliyan[2], Camilo Obando[1]

[1] Johnson & Johnson, Titusville, NJ, USA
[2] OMNY Health., Atlanta, Georgia, USA

Background

- Bipolar depression is a chronic, severe, and debilitating psychiatric condition characterized by recurrent depressive episodes that significantly impair daily functioning and quality of life.[1]
- Managing bipolar depression remains clinically challenging, as many patients experience persistent symptoms despite initial pharmacological interventions.[2]
- Lumateperone—an atypical antipsychotic with a unique multimodal mechanism of action acting upon serotonergic, dopaminergic, and glutamatergic pathways—is an FDA-approved treatment for bipolar depression.[3]
- While clinical trials have demonstrated efficacy, real-world evidence regarding longitudinal depressive symptom trajectory following lumateperone initiation in routine clinical practice remains limited.

Objective

To evaluate real-world effectiveness and longitudinal change in depressive symptom severity, measured by the Patient Health Questionnaire-9 (PHQ-9), over 3 to 6 months among adult patients with bipolar depression initiating lumateperone treatment in the United States.

Methods

Study design and data sources

- A retrospective observational cohort study was conducted using de-identified electronic health records (EHR) from the nationwide OMNY Health platform and linked open claims between January 1, 2022, and June 1, 2026.
- The index date was defined as the date of the earliest documented lumateperone prescription on or after January 1, 2022.

Study population (Figure 1)

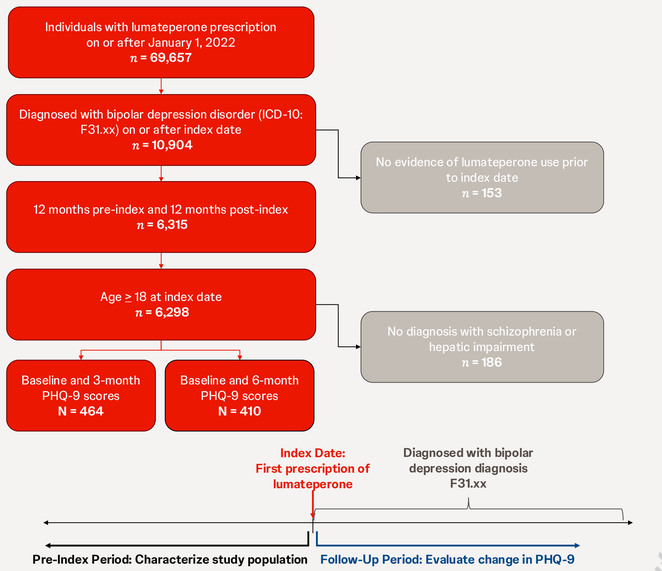
Inclusion Criteria:

- Adults (≥ 18 years at index) with ≥1 prescription for lumateperone.
- Diagnosed with bipolar depression (ICD-10-CM code: F31.xx) on or after the index date.
- Documented baseline PHQ-9 score measured within 30 days prior to or up to 7 days post-index date.
- Continuous healthcare touchpoints (12 months pre-index and 12 months post-index).

Exclusion Criteria:

- Prior evidence of lumateperone use during the pre-index period.
- Diagnosis of schizophrenia or hepatic impairment during the study observation window.

Figure 1. Patient Selection & Study Design



Outcomes and analyses

Primary Outcome:

- Change in total PHQ-9 score from baseline to Month 3 (+46 to +135 days) and Month 6 (+136 to +225 days).

Secondary & Exploratory Outcomes:

- Time to achieve clinically meaningful improvement, defined as a ≥5-point reduction in total PHQ-9 score from baseline.
- Model-adjusted longitudinal changes in PHQ-9 overall and stratified across baseline severity tiers (None/Minimal [0–4], Mild [5–9], Moderate [10–14], Moderately Severe [15–19], Severe [20–27]).

Statistical Analysis:

- Paired two-sided hypothesis tests (t-tests and Wilcoxon signed-rank) assessed unadjusted changes from baseline to Month 3 and Month 6.
- Cumulative Incidence: Non-parametric cumulative incidence functions (CIF) evaluated the time-to-≥ 5-point PHQ-9 improvement, accounting for censoring.
- Inverse Probability Weighting (IPW): Visit Intensity Weighting (via Cox proportional hazards modeling) and Measurement Weighting (via generalized linear modeling) were combined to adjust for informative drop-out and irregular EHR visit frequencies.
- Longitudinal Modeling: Weighted piecewise linear mixed-effects models (LMM) with random intercepts were fitted to estimate marginal adjusted PHQ-9 trajectory and changes across Months 3, 6, 9, and 12.

Key Findings

Baseline Characteristics of Individuals with PHQ-9 at month 3 (n =464)

- **Demographics:** Mean age was 45 years (92% < 65 years); 77% were female. Geographic distribution was predominantly Midwest (53.0%) and South (45.6%).
- **Clinical History:** High psychiatric comorbidity was observed, including anxiety disorders (93%), depression (86%), and ADHD (35%).
- **Prior Treatments:** Pre-index medication use was substantial: antidepressants (98%), antipsychotics (69%), benzodiazepines (63%), anticonvulsants (54%), and lithium (16%). Evidence-based psychotherapy history was present in 55%.
- **Baseline Severity:** Mean baseline PHQ-9 was 10.6 (SD 8.1) overall. Baseline severity breakdown: 34% None/Minimal, 13% Mild, 18% Moderate, 18% Moderately Severe, and 17% Severe.
- Similar trends were observed among the subset of patients with PHQ-9 at month 6.

Table 1. Baseline demographic and clinical characteristics)

	PHQ-9 at month 3 n = 464	PHQ-9 at month 6 n = 410
Baseline characteristics		
Age, mean(SD)	45.1 (13.4)	45.4 (13.1)
Age < 65, n(%)	425 (91.6%)	374 (91.2%)
Female, n(%)	355 (76.5%)	317 (77.3%)
Baseline PHQ-9, mean(SD)	10.6 (8.1)	10.9 (8.0)
Baseline comorbidities		
ADHD	162 (34.9%)	133 (32.4%)
Anxiety	433 (93.3%)	373 (91.0%)
Depression	397 (85.6%)	344 (83.9%)
Hyperlipidemia	182 (39.2%)	163 (39.8%)
Hypertension (high blood pressure)	231 (49.8%)	219 (53.4%)
Prior treatments		
Anticonvulsants	252 (54.3%)	207 (50.5%)
Antidepressants	456 (98.3%)	397 (96.8%)
Atypical antipsychotics	321 (69.2%)	292 (71.2%)

Unadjusted Mean PHQ-9 Score Reduction

- 3-Month Follow-Up (n=464): Mean PHQ-9 score decreased by -2.74 points (SD 7.93; median change: -1.0; p<0.001). significantly
- 6-Month Follow-Up (n=410): Mean PHQ-9 score notably decreased by -3.48 points (SD 8.41; median change: -2.0; p<0.001).

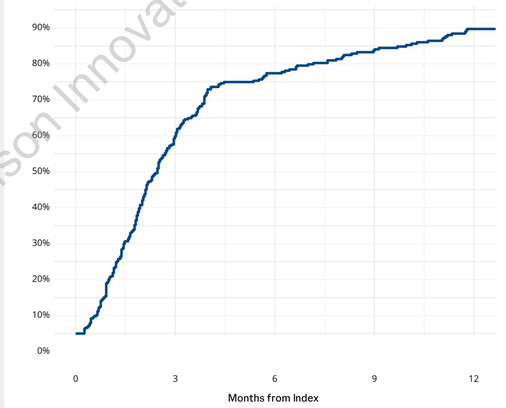
Cumulative Incidence of Clinical Improvement (≥ 5-point drop)

- The probabilities of achieving a ≥5-point improvement in PHQ-9 score at Months 3, 6, 9, and 12 were 55%, 72%, 79%, and 85%, respectively **(Figure 2)**.

Model-Adjusted Longitudinal Outcomes (IPW-Weighted LMM)

- Overall Adjusted Changes: Meaningful reductions in adjusted mean PHQ-9 were sustained across all follow-up intervals (**Figure 3**).
- Among baseline severity subgroups, patients entering with higher symptom burden achieved the greatest absolute score reductions at Month 6 (**Table 2**).

Figure 2. Cumulative Incidence of PHQ-9 Clinical Improvement



	Risk Table				
	Month 0	Month 3	Month 6	Month 9	Month 12
At Risk	308	135	79	54	34
Events	0	168	220	237	252

	Cumulative Incidence (95% CI)			
	Month 3	Month 6	Month 9	Month 12
All Patients	55% (49%-60%)	72% (67%-77%)	79% (73%-83%)	85% (80%-88%)

Figure 3. Model-Predicted PHQ-9 Trajectories Across 12 months
Stratified by Baseline Severity (IPW-Adjusted Mixed-Effects Model)

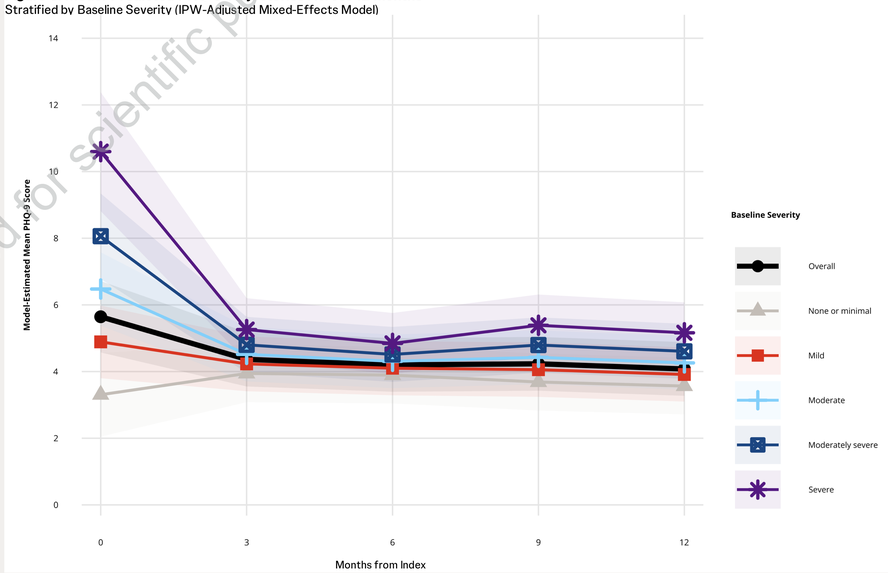


Table 2. Change in PHQ-9 score from index
(IPW Model estimates)

Baseline severity	Month 3	Month 6	Month 9	Month 12
Overall	-1.3 (-2.1, -0.5)	-1.4 (-2.2, -0.6)	-1.4 (-2.2, -0.6)	-1.6 (-2.4, -0.8)
None or minimal	0.6 (-0.2, 1.5)	0.6 (-0.3, 1.4)	0.4 (-0.5, 1.2)	0.3 (-0.6, 1.1)
Mild	-0.7 (-1.5, 0.2)	-0.8 (-1.6, 0.0)	-0.8 (-1.6, -0.0)	-1.0 (-1.8, -0.2)
Moderate	-2.0 (-2.8, -1.1)	-2.2 (-3.0, -1.4)	-2.0 (-2.9, -1.2)	-2.2 (-3.0, -1.4)
Moderately severe	-3.3 (-4.1, -2.4)	-3.5 (-4.4, -2.7)	-3.3 (-4.1, -2.4)	-3.5 (-4.3, -2.6)
Severe	-5.3 (-6.3, -4.4)	-5.8 (-6.7, -4.8)	-5.2 (-6.1, -4.3)	-5.4 (-6.3, -4.5)

Conclusions

- In this nationwide real-world EHR study, initiation of lumateperone was associated with statistically significant and sustained reductions in depressive symptom severity among patients with bipolar depression.
- Clinically meaningful score reductions (≥ 5-point PHQ-9 improvement) were achieved by 43% of patients at Month 6 and 55% by Month 12.
- Patients with higher baseline depression severity experienced the most pronounced symptom gains, driving overall cohort improvements into sub-threshold severity ranges.
- These findings complement clinical trial data by demonstrating the real-world effectiveness of lumateperone in routine psychiatric care.

Limitations

- Observational Data: Reliance on EHR structured data may lead to residual confounding, unmeasured variables (e.g., treatment adherence, socioeconomic status), or missing non-EHR clinical encounters.
- Informative Measurement: PHQ-9 documentation occurs during routine office visits rather than fixed protocol intervals; although IPW modeling was utilized to mitigate visit intensity bias, residual sampling selection bias remains possible.
- Single-Arm Design: Lack of a direct active comparator group limits direct comparative effectiveness inferences against other atypical antipsychotics in this specific dataset.

Acknowledgments

This study was funded by Johnson & Johnson Innovative Medicine.

Disclosures

- MN, and CO are employees of Johnson & Johnson Innovative Medicine

Mood Disorder

Presented at: 2026 Psych Congress, September 15-19, 2026. New Orleans, Louisiana | USA

References:

1. McIntyre RS, Berk M, Brietzke E, Goldstein BI, López-Jaramillo C, Kessing LV, et al. Bipolar disorders. Lancet Lond Engl. 2020 Dec;396(10265):1841–56. doi:10.1016/S0140-6736(20)31544-0
2. Sajatovic M. Bipolar disorder: disease burden. Am J Manag Care. 2005 Jun;11(3 Suppl):S80-4.
3. Cooper D, Gupta V. Lumateperone. StatPearls [Internet]. 2023. Available from: https://www.ncbi.nlm.nih.gov/books/NBK560844/