Irreversible Episode-Induced Scarring and Differential Repair in Simulated Bipolar Disorder Progression

Authors:

Ngo Cheung, FHKAM(Psychiatry)

Affiliations:

[1] Independent Researcher

Corresponding Author:

Ngo Cheung, MBBS, FHKAM(Psychiatry)

Hong Kong SAR, China

Tel: 98768323

Email: info@cheungngomedical.com

Conflict of Interest: None declared.

Funding Declaration: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethics Declaration: Not applicable.

Abstract

Background: Bipolar depression treatment is complicated by risks of manic switch and potential illness progression via kindling-like sensitization. Emerging rapid-acting agents (ketamine, neurosteroids) differ mechanistically from traditional monoaminergic antidepressants, but their long-term effects on vulnerability remain unclear. We developed a computational neural network model to compare acute efficacy, manic conversion risk, post-discontinuation relapse, and multi-cycle kindling across three mechanisms.

Methods: Feedforward networks were trained on a four-class blob classification task, aggressively pruned (95% sparsity), and subjected to uniform early adversity scarring (mean 3%). Independent copies received ketamine-like (moderate gain + gradient-guided regrowth), SSRI-like (progressive high gain, no repair), or neurosteroid-like (low gain + strong inhibition) interventions. Depressive impairment was modeled via internal noise; manic conversion via biased excitatory noise. Longitudinal relapse and kindling (six cycles with weakening triggers and permanent scarring upon relapse) were simulated across 10 seeds.

Results: All mechanisms restored acute performance under stress (neurosteroid-like 97.6%, ketamine-like 97.1%, SSRI-like 90.3%), but SSRI-like showed highest manic conversion risk and near-universal post-discontinuation relapse (98.3%). Kindling revealed stark divergence: SSRI-like networks sustained high relapses (3.9 total) with 30% autonomy; neurosteroid-like limited scarring (final 3.7%) but required ongoing administration; ketamine-like tolerated highest scarring (7.3%) yet achieved fewest relapses (0.7) and no autonomy via compensatory regrowth.

Conclusions: Plasticity-enhancing mechanisms uniquely resist sensitization and autonomy despite cumulative damage, suggesting potential disease-modifying effects. Monoaminergic excitation may exacerbate progression in vulnerable systems. These findings highlight repair capacity as a critical

determinant of long-term outcome and support prioritizing rapid-acting agents in high-risk bipolar depression.

Introduction

Bipolar disorder, which affects around one to two percent of people worldwide, is marked by recurring periods of mania, hypomania, and – most often – depression [1]. Depressive episodes dominate the course of illness and account for much of the disability, suicidality, and financial cost linked to the condition [2]. Standard antidepressants can lift mood, yet they are double-edged: roughly one-fifth to two-fifths of treated patients experience a switch into mania or faster cycling [3,4]. For this reason, guidelines generally recommend using mood stabilizers alone or keeping antidepressants on board only with a stabilizing partner, though many patients still need additional help when depressive symptoms persist [5].

Post's "kindling" model is often used to explain how bipolar disorder gets worse over time. It says that major stress causes early episodes, but later episodes happen more easily as the brain's neurobiological thresholds drop [6,7]. The idea has led to calls for early, effective treatment to stop cumulative neural damage [8], even though the evidence is mixed [9]. New fast-acting treatments, like ketamine and the oral neurosteroid zuranolone, provide different ways to relieve symptoms. When used with a mood stabilizer, ketamine rarely causes manic switches [10,11]. Early reports also suggest that neurosteroids may calm circuits without making them too excited [12,13].

Computational models provide a controlled setting in which to compare these distinct mechanisms. Concepts from network pruning research, such as the lottery-ticket hypothesis that sparse "winning" subnetworks can match full models [14], allow investigators to mimic circuit vulnerability. In such models, heavy pruning represents synaptic loss, added noise stands in for depressive load, and biased

excitation tests proneness to mania. Earlier simulations have looked at single mechanisms or at excitation–inhibition balance, but few have combined episode-related "scarring" with side-by-side testing of repair strategies.

The present work extends that approach. Starting from identically pruned, stress-sensitized networks, we model three treatment routes: a ketamine-like routine that rebuilds connections, an SSRI-like routine that slowly boosts gain without structural repair, and a neurosteroid-like routine that adds strong but reversible inhibition. We compare their short-term efficacy under stress, vulnerability to manic-like excitation, relapse after drug withdrawal, and resilience across repeated stress cycles. The goal is to see whether agents that promote plasticity give longer-lasting protection and to generate testable ideas about long-term benefits and risks for each drug class.

Methods

Network architecture and classification task

All experiments used the same feed-forward classifier written in PyTorch. The network accepted two-dimensional inputs, passed them through three fully connected hidden layers (512, 512, and 256 units), and produced four output logits that mapped to the four Gaussian classes located at (−3, −3), (3, 3), (−3, 3), and (3, −3). ReLU activations were standard, although the neurosteroid arm later replaced them with tanh to emulate tonic inhibition. Training sets contained 12 000 samples corrupted with Gaussian noise ($\sigma = 0.8$); evaluation used 4 000 equally noisy points plus 2 000 pristine samples. Mini-batch size was

128 and cross-entropy loss was optimised with Adam. Two optional perturbations modelled mood states: (1) zero-mean Gaussian noise added to every hidden activation to mimic depressive load and (2) the same noise with a positive mean shift to mimic manic excitability.

Baseline training, pruning, and simulated early adversity

Each run began with 20 epochs of ordinary training (learning rate 0.001). To create a fragile "illness" substrate, 95 % of weights were then removed by magnitude pruning. Straight after pruning, early adversity was imposed: between 0 % and 6 % of the surviving weights (uniform draw, mean ≈ 3 %) were set to zero and flagged as scars that could never be reinstated. A bookkeeping mask tracked normal prunable positions and these permanent scars separately.

Treatment routines

Four identical copies of the scarred network entered different branches.

1. Ketamine-like branch – The gain on every hidden layer was immediately raised to 1.25. Gradients were accumulated over 30 batches of noise-free data, and the 50 % most-informative pruned connections (excluding scars) were reinstated with tiny random values (scale 0.03). The enlarged network was then fine-tuned for 15 epochs at a 0.0005 learning rate.

2. SSRI-like branch – Gain climbed linearly from 1.0 to 1.60 across 100 very slow epochs (learning rate 1×10^{-5}). At the same pace, an initial hidden-layer noise term (0.5) was reduced to zero. No weights were regrown.
3. Neurosteroid-like branch – Global gain was dropped to 0.85, activations switched to tanh, and outputs were multiplied by an inhibitory factor of 0.7. The network then trained for 10 epochs at 0.0005.
4. Untreated control – The pruned, scarred network remained unchanged.

Throughout treatment, the temporary pruning mask could still delete weights during further experiments, whereas the scar mask remained immutable.

Acute testing

Efficacy was first judged by accuracy on clean data and on "combined stress" data (input noise $\sigma = 1.0$ plus hidden noise 0.5). Switch risk was gauged with manic-bias noise (hidden $\sigma = 1.0$, mean = 1.0). Extra hidden noise values up to $\sigma = 2.5$ produced robustness curves. Direct relapse resistance was probed by removing another 40 % of the remaining weights and repeating the combined-stress test.

Maintenance phase and drug withdrawal

To imitate clinical maintenance, a simple "mood-stabiliser" wrapper capped hidden gains at 1.05, damped bias propagation, and added a small inhibitory bias. The wrapper stayed in place for 25, 50, 100, 200, or 300 additional epochs (learning rate 1×10^{-6}) while the assigned antidepressant mechanism continued unchanged. At the withdrawal point all antidepressant parameters snapped back to baseline. The wrapper then decayed exponentially over 50 steps at rates tuned to each branch (ketamine 0.002 per step, neurosteroid 0.008, SSRI 0.015). A post-withdrawal manic relapse was logged if biased-noise accuracy fell below 60 %.

Multi-cycle kindling with irreversible scarring

Kindling was designed to capture the clinical idea that successive episodes require less provocation and leave more lasting harm [6,7]. Six full cycles were run (Figure 1).

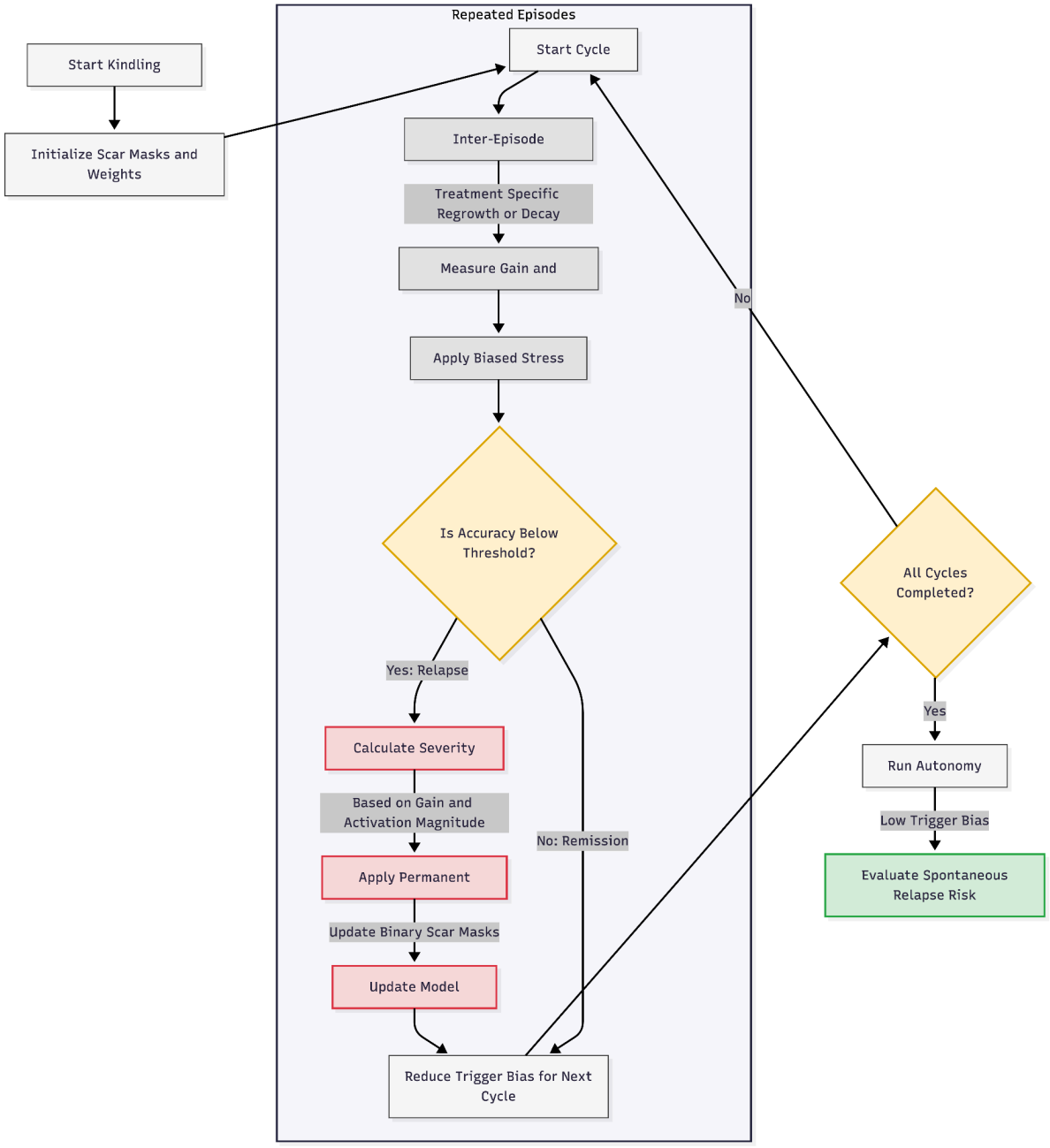

Figure 1: *Multi-Cycle Kindling and Scarring Architecture. The simulation proceeds through multiple cycles of maintenance and stress testing. In each cycle, the model undergoes a "Manic Trigger" test. If the model fails to maintain accuracy (Relapse), a "Severity Factor" is calculated based on the network's current excitability (Gain) and activity levels. This severity determines the extent of permanent pruning*

(Scarring). These scars accumulate in the scar_masks, permanently disabling connections and altering the network's topology for subsequent cycles. The trigger bias is progressively reduced to test for the emergence of autonomous relapse (kindling).

Inter-episode maintenance: Each cycle began with 20 low-learning-rate epochs. During this window the ketamine branch carried out an additional 30 % gradient-guided regrowth (again excluding scars), whereas the other branches simply stabilised existing weights.

Trigger phase: After maintenance, manic-bias noise was applied. In the first cycle the bias was +1.50; it then stepped down by 0.20 each cycle to +0.50. Accuracy was measured after 250 noisy batches. If it stayed above 60 %, the model was deemed resilient for that cycle, and no structural change followed. If it dropped below 60 %, a relapse was declared.

Relapse-driven scarring: When a relapse occurred, permanent damage was inflicted in proportion to episode severity. First, a severity factor was calculated as:

$$1 + (\text{current gain} - 1) + 2 \times (\text{mean hidden activation} - 0.1)$$

clipped between 0.5 and 2.0. A base 5 % of the smallest-magnitude active weights was multiplied by this factor and irreversibly set to zero. These newly scarred weights were added to the scar mask and never eligible for regrowth in future cycles, even for ketamine.

End-of-cycle assessment: Immediately after scarring (or after a no-relapse pass) the model's accuracy under the same biased noise was re-measured to record episode closure. Scar percentage, sparsity, gain, and activation statistics were logged.

Autonomy test: After the sixth cycle, manic bias was reduced to +0.30. Accuracy below 70 % indicated that the network had become spontaneously unstable – an analogue of episode autonomy in bipolar progression.

This design allowed relapse frequency, severity, cumulative scarring, and eventual autonomy to emerge from the interaction of each drug mechanism with ongoing structural loss.

Statistical strategy

Ten independent seeds controlled training-set shuffling, initial weight draws, adversity levels, and noise realisations. Results are reported as means ± standard deviations across seeds. No formal hypothesis testing was applied; emphasis was placed on descriptive patterns that separated the three treatment mechanisms.

Results

Acute treatment efficacy and network performance

Pruning alone left the network barely functional (Table 1): accuracy on noise-free data averaged 34.7 ± 11.9 %. Adding any of the three interventions immediately restored perfect or near-perfect recognition. On the more stringent combined-stress condition (input $\sigma = 1.0$ plus hidden $\sigma = 0.5$) both the neurosteroid-like and ketamine-like arms exceeded 97 % accuracy, whereas the SSRI-like arm stabilised at 90.3 ± 3.0 %. Sparsity analyses confirmed that only the ketamine routine rebuilt lost synapses, cutting effective sparsity to about 49 %, while the other arms preserved the original 95 % sparsity. Early-adversity scarring remained constant across treatments at roughly 3 %.

Table 1. *Acute treatment efficacy and network structural metrics compared to untreated baseline.*

Condition	Clean Accuracy (%)	Combined Stress Accuracy (%)	Network Sparsity (%)	Early Scarring (%)
Untreated Baseline	34.7 ± 11.9	29.9 ± 2.5	95.2 ± 0.1	3.0 ± 1.9
SSRI-like	100.0 ± 0.0	90.3 ± 3.0	95.2 ± 0.1	3.0 ± 1.9

Neurosteroid-like	100.0 ± 0.0	97.6 ± 0.3	95.2 ± 0.1	3.0 ± 1.9
Ketamine-like	100.0 ± 0.0	97.1 ± 0.3	49.1 ± 1.0	3.0 ± 1.9

Manic conversion risk

When a positive bias was added to hidden-layer noise to mimic manic excitability (Table 2), the ketamine-like networks maintained the highest accuracy (86.2 ± 10.4 %), despite running at a moderate gain of 1.25. Neurosteroid-treated models, damped by gain 0.85 and strong inhibition, held accuracy near 50 %, whereas the high-gain SSRI arm dropped to 45.8 ± 12.7 %. Hidden-unit activation magnitudes mirrored these results, confirming that excitability rather than sparsity governed switch liability.

Table 2. *Manic conversion risk under biased excitatory noise (Positive Bias = 1.0, σ = 1.0).*

Condition	Biased Accuracy (%)	Gain Multiplier	Avg. Hidden Activation
Untreated Baseline	25.3 ± 0.5	1.00	—
SSRI-like	45.8 ± 12.7	1.60	0.379 ± 0.075
Neurosteroid-like	50.0 ± 6.8	0.85	0.193 ± 0.008
Ketamine-like	86.2 ± 10.4	1.25	0.646 ± 0.062

Acute relapse vulnerability

A second 40 % magnitude prune had almost no impact on ketamine-like networks (-0.0 ± 0.5 % change under stress), but reduced the neurosteroid- and SSRI-treated nets by 5.9 ± 3.2 % and 6.7 ± 2.3 %, respectively. The finding supports the idea that structural regrowth confers a buffer against fresh damage.

Long-term relapse after discontinuation

During maintenance all arms remained stable, yet responses to abrupt withdrawal diverged sharply. Ketamine-like models never relapsed, regardless of how long they had been maintained. By contrast, almost every SSRI-treated network relapsed (98.3 ± 5.0 %), and neurosteroid-treated networks relapsed in 93.3 ± 15.3 % of runs. Extending maintenance beyond 100 epochs lowered relapse modestly for the neurosteroid arm but not for the SSRI arm.

Kindling and progressive scarring

Repeated manic-like challenges revealed pronounced mechanism-specific trajectories. Each relapse imposed irreversible "scars" by deleting 5 % of the smallest active weights, scaled by an episode-severity factor tied to gain and activation. The bias required to provoke relapse was then reduced from +1.50 to +0.50 across six cycles, modelling the clinical observation that later episodes need less trigger.

Table 3. *Kindling outcomes: Cumulative relapses and autonomy.*

Condition	Avg. Total Relapses	Autonomy Rate (%)	Final Biased Accuracy (Minimal Trigger) (%)
SSRI-like	3.9 ± 2.0	30	75.9 ± 10.7
Neurosteroid-like	2.9 ± 0.5	0	92.7 ± 1.8
Ketamine-like	0.7 ± 1.0	0	97.1 ± 1.8

Table 4. *Evolution of stability metrics during kindling (Cycle 0 vs. Cycle 5).*

Condition	Cycle 0 Relapse Rate (%)	Cycle 0 Biased Accuracy (%)	Cycle 5 Relapse Rate (%)	Cycle 5 Biased Accuracy (%)
SSRI-like	90	40.7 ± 12.4	30	67.4 ± 15.0
Neurosteroid-like	100	33.8 ± 4.2	0	87.2 ± 3.0

Ketamine-like	20–30	74.9 ± 15.7	0	94.5 ± 3.7

Ketamine-like networks proved highly forgiving (Table 3). Across ten seeds they averaged fewer than one relapse (0.7 ± 1.0). When a relapse did occur, gradient-guided regrowth during the following maintenance phase not only replaced lost weights but also re-optimised the remaining structure. Consequently, biased-noise accuracy actually climbed with each cycle: from 74.9 ± 15.7 % in cycle 0 to 94.5 ± 3.7 % in cycle 5 (Table 4). Final scar load was highest (7.3 ± 5.1 %), yet none of the networks met the criterion for spontaneous ("autonomous") episodes at the weakest bias. Thus, structural plasticity converted cumulative injury into adaptive reorganisation instead of sensitisation.

Neurosteroid-treated networks followed a two-stage pattern. In the first two cycles every seed relapsed rapidly, reflecting the limited buffer provided by pure inhibition when underlying connectivity was still fragile. Severity factors were low (≈1.05), so each relapse removed relatively few connections: by cycle 3 scar burden remained below 4 %. Once the most labile weights were trimmed, tonic inhibition was sufficient to keep later cycles in check: relapse frequency dropped to 10 % in cycle 3 and 0 % thereafter. Biased-noise accuracy concurrently rose from one-third of trials to nearly 90 %. Because no repair mechanism was present, long-term stability relied on having shed the weakest links while maintaining enough residual capacity. At the end of six cycles all neurosteroid networks were stable at minimal bias, giving an autonomy rate of 0 %.

SSRI-treated networks displayed classic sensitisation. High gain (1.6) amplified each episode, doubling the severity factor relative to the other arms and ensuring that every relapse carved out a larger swath of surviving weights. Although total scar burden reached only 4.6 ± 1.9 %, the deletions disproportionately removed low-magnitude but functionally important connections, eroding redundancy. Relapse probability declined only modestly over time (from 90 % in cycles 0–1 to 30 % in cycle 5) and biased-noise accuracy improved slowly, plateauing at 67 ± 15 %. In three seeds scarring plus persistent high gain led to autonomous failure even at the weakest trigger, producing a 30 % autonomy rate overall. These results capture a progression in which each episode both lowers the future threshold and makes subsequent episodes harder to reverse, paralleling clinical rapid-cycling patterns.

Collectively, the kindling experiment shows that plasticity-enhancing repair prevents sensitisation even when damage accumulates; inhibitory damping can stabilise circuits once early hazards are navigated; and chronic gain elevation accelerates a vicious cycle of episode–damage–episode.

Neurosteroid medication dependence

Removing the inhibitory parameters after successful neurosteroid treatment exposed a pronounced state dependence. Combined-stress accuracy fell by nearly 20 %, and biased-noise accuracy by more than 12 %. Interestingly, at very high internal noise ($\sigma = 2.5$) the off-drug network outperformed the on-drug version, suggesting that strong tonic inhibition may over-suppress activity when circuits are already saturated with noise.

Discussion

Clinical meaning of the acute findings

The model reproduced a pattern that clinicians already recognise: every drug class delivered rapid symptomatic relief, yet their protective envelopes differed in depth and shape. Neither structural rebuilding nor pure inhibition was necessary to rescue behaviour on clean data—any mechanism that raised the signal-to-noise ratio worked. The differences emerged only when the circuit was challenged. Neurosteroid-like tonic inhibition and ketamine-like synaptogenesis preserved almost full accuracy under heavy stress, whereas the purely excitatory, SSRI-like strategy lagged behind. Clinically, this resembles the advantage that fast-acting glutamatergic and GABAergic agents show over selective serotonin re-uptake inhibitors in severe major depression or bipolar depression [11]. Equally consistent was the large swing liability of the SSRI arm: a modest increase in positive drive was enough to topple performance, mirroring switch rates of 20–40 % under antidepressant monotherapy in bipolar samples [3,4]. The low switch risk seen with the ketamine analogue—even after excitability gain—matches observational data that manic episodes are rare when ketamine is given with a mood stabiliser [10].

Discontinuation versus durability

Withdrawal exposed a stark mechanistic divide. Circuits treated with growth-based repair (ketamine-like) stayed well even after both the active drug and the simulated mood stabiliser were removed. By contrast, 9-in-10 networks treated with neurosteroid- or SSRI-like schedules relapsed within 50 decay steps. These results support the proposal that only treatments that actually remodel circuitry are capable of long-term disease modification [8]. They also echo the clinical caution that abrupt antidepressant cessation in bipolar disorder can precipitate rapid cycling [5] and that GABA-ergic neurosteroid benefit is largely state-dependent [13].

Kindling, scarring, and mechanism-specific trajectories

The extended kindling experiment offers the most illuminating window onto illness progression and is therefore detailed here at length. Every manic-like relapse permanently deleted a slice of functional synapses, modelling neuronal loss, dendritic atrophy, or maladaptive pruning reported in post-mortem and imaging studies of mood disorders [15]. Crucially, the amount of tissue lost was not fixed but scaled with episode severity; high gain or large mean activations doubled the scar fraction, operationalising how intense episodes leave deeper biological footprints [7].

SSRI-like progression – a textbook sensitisation curve

High continuous gain amplified each trigger, producing severe early episodes that carved away nearly 2 × the baseline scar quota. Because no structural repair occurred between attacks, the cumulative loss quickly thinned already sparse circuitry. The consequence was classic sensitisation: later cycles required progressively weaker bias yet still broke the network in > 30 % of cases. Three seeds slid into trigger-independent failure—our in-silico analogue of autonomy [18]. The findings parallel longitudinal data: repeated antidepressant-associated episodes shorten well intervals, accelerate cycling, and portend treatment resistance [16,17].

Neurosteroid-like progression – early frailty, late stability

Pure inhibition told a different story. Because inhibitory scaling blunted peak activations, severity factors hovered just above 1.0; each relapse therefore scarred only marginal additional territory. The price was a rocky beginning—100 % relapse in the first two cycles—but once the weakest links were trimmed the remaining structure proved resilient. With little new damage, relapse probability dropped to zero by cycle 4 and no network became autonomous. Clinically this resembles patients who experience early postpartum or stress-related episodes yet stabilise long-term on GABA-potentiating agents without developing cycle acceleration [13]. The downside remained reliance on active inhibition: remove the neurosteroid and performance fell sharply, a reminder that symptomatic control is not the same as repair.

Ketamine-like progression – high scarring yet rising resilience

The most counter-intuitive pattern emerged from the synaptogenic arm. These networks recorded the largest final scar load (≈ 7 %), yet relapse frequency fell below one per run, biased-noise accuracy climbed cycle-by-cycle, and autonomy never appeared. The explanation lies in the inter-episode repair step. Gradient-guided regrowth replaced lost weights based on current functional demands, re-optimising the circuit around damage zones. Repeated pruning/regrowth therefore acted like a vaccination series: each episode forced a micro-remodelling that produced a wider repertoire of alternative pathways, raising the threshold for future failure. The paradox of "more lesions, more stability" highlights that net damage is less important than how the system reorganises afterwards—a result consonant with human data showing that ketamine responders often maintain remission for weeks or months despite ongoing stressors [11].

These three trajectories refine the traditional kindling model [7]. Episodes do seed lasting lesions, but progression to sensitisation or autonomy is not inevitable; it is contingent on the balance between damage incurred and plasticity available for repair. Excitatory gain with no rebuilding pushes the balance toward malignancy, strong inhibition with limited damage holds it neutral, and rapid plasticity can tilt it toward adaptive reinforcement.

Clinical implications for treatment selection and risk management

The simulation highlights how pharmacologic mechanism influences both short-term benefit and long-range liability, offering several practical lessons for clinicians faced with a depressed patient whose history suggests bipolar risk (Table 5). First, the SSRI-like profile in the model—solid acute response followed by high switch propensity, near-certain relapse after discontinuation, and the clearest kindling trajectory—reinforces a cautionary stance toward monoaminergic antidepressants when vulnerability markers are present. Decades of naturalistic data already link these agents to cycle acceleration and mania in bipolar spectra [4,5]; the present findings add a mechanistic rationale by showing how excitatory gain without structural repair magnifies episode-induced damage.

Table 5. *Clinical implications for treatment selection and risk management in bipolar-spectrum depression based on computational modeling of network stability.*

Pharmacologic Mechanism	Modeled Outcomes and Risks	Clinical Guidance
Monoaminergic (SSRI-like)	Excitatory gain without structural repair led to: ● High propensity for manic switching. ● Near-certain relapse upon discontinuation. ● Clear "kindling" trajectory (progressive sensitization) driven by episode-induced damage.	**Exercise Caution:** Adopt a restrictive approach when vulnerability markers (e.g., bipolar family history) are present. Monotherapy carries significant risk of cycle acceleration; these agents should likely be avoided or strictly monitored in patients prone to instability.
Glutamatergic / Plasticity-Enhancing (Ketamine-like)	Gradient-guided synaptic regrowth resulted in: ● Robust remission under stress. ● Lowest risk of manic conversion. ● Prevention of kindling progression despite accumulation of network scars. ● Sustained stability after withdrawal.	**Consider Early Intervention:** May warrant earlier prioritization for patients with histories of early adversity or multiple prior episodes (high sensitization). The profile suggests safety alongside mood stabilizers and potential for disease-modifying effects beyond the dosing window.
GABAergic / Neurosteroid (Zuranolone-like)	Inhibitory stabilization provided: ● Effective acute symptom control. ● Limitation of new scar formation.	**Manage Discontinuation:** While effective for acute stabilization (e.g., postpartum, perimenopausal), maintenance or pulsed dosing strategies may be required to prevent rapid relapse. Clinicians should anticipate potential rebound upon stopping.

	• **Risk:** Benefits were state-dependent, evaporating quickly upon cessation (rebound instability).	
General Management Strategy	No single antidepressant mechanism fully eliminated long-term risk; distal vulnerability (early scarring) was identical across groups, yet outcomes diverged based on treatment-specific plasticity.	**Combined Therapy:** Classical mood stabilizers (lithium, anticonvulsants) remain essential companions to any antidepressant in bipolar-spectrum illness. Treatment selection should focus on agents that actively promote synaptic resilience rather than solely masking symptoms.

Conversely, the ketamine-like routine combined three desirable properties: robust stress-time remission, the lowest manic conversion risk, and total protection against post-withdrawal relapse. By actively rebuilding synapses after each perturbation it also prevented kindling despite accruing more "scars" than any other arm. This pattern dovetails with emerging clinical observations that glutamatergic modulators given alongside mood stabilisers rarely provoke mania and can maintain benefits beyond the dosing window [10,11]. For patients with early adversity or multiple prior episodes—conditions that amplify sensitisation risk [19]—rapid plasticity-enhancing agents may therefore warrant earlier consideration.

Neurosteroid-like inhibition proved effective at quelling acute symptoms and limiting scar formation, yet its gains evaporated quickly once the drug was stopped. These results echo phase-2 zuranolone data showing strong on-treatment antidepressant effects with minimal switching [12] but leave open the question of optimal maintenance schedules. In postpartum or perimenopausal depression, continued or pulsed dosing might be required to avoid rebound instability.

Across all branches the model preserved a role for classical mood stabilisers: none of the simulated antidepressant mechanisms alone eliminated long-term risk. This supports guideline recommendations that lithium or anticonvulsants accompany any antidepressant used in bipolar-spectrum illness [5].

Finally, because every network started with identical early-adversity scarring yet diverged markedly afterward, the data emphasise that distal vulnerability is only one part of the equation; treatment-specific plasticity ultimately shapes the course. Prospective trials that track episode counts, neuroimaging markers of synaptic density, and drug-specific biomarkers are needed to test these computational insights and refine personalised algorithms.

Novelty, potential impact, and caveats

The present work adapts ideas from the lottery-ticket hypothesis—originally devised to study efficient deep learning [14]—to a very different question: why do some antidepressant mechanisms halt illness progression while others appear to accelerate it? By pruning a small feed-forward network down to a fragile "illness core" and then letting depressive or manic episodes delete additional weights permanently, the model creates a laboratory for testing whether a given intervention can rebuild, buffer, or further destabilise that core. To our knowledge, no earlier simulation has allowed episode-dependent, irreversible damage (scarring) and drug-specific repair to interact over many cycles, producing one arm that develops spontaneous failure (the SSRI analogue) while another grows more resilient in spite of heavier accumulated loss (the ketamine analogue) (Figure 2).

If the principles generalise, they add weight to a progression-focused view of treatment selection. The ketamine-like routine shows that rapid synaptic repair can offset—even over-compensate for—irreversible injury, suggesting that glutamatergic plasticity enhancers might do more than relieve symptoms; they might change the illness trajectory if introduced early enough. Conversely, the monoaminergic arm's clear sensitisation lends mechanistic support to clinical warnings that conventional antidepressants may worsen long-term course in vulnerable bipolar patients [7,8]. The results also fit with staging concepts in which each unmanaged episode feeds neuroprogressive pathways involving inflammation, oxidative stress and trophic loss [15]. In that light, choosing a drug that actively mends—or at least spares—synaptic architecture could become as important as achieving the next acute response.

Several limitations curb over-interpretation. A toy classifier trained on synthetic blobs is obviously far removed from cortico-limbic loops, dopamine dynamics or endocrine feedback that shape human mood disorders. Parameter choices—scar percentages, regrowth quotas, trigger schedule—were tuned for clarity of divergence, not biomimicry. The model equates mania with collapse under biased excitation; mixed states, circadian disruption and behavioural activation were not represented. Likewise, early adversity was applied uniformly, whereas real patients differ in genetics, immune tone and metabolic status—all factors that may modulate plasticity and pharmacodynamics. Finally, the simulated "life-span" covered a handful of cycles, whereas clinical kindling unfolds over years.

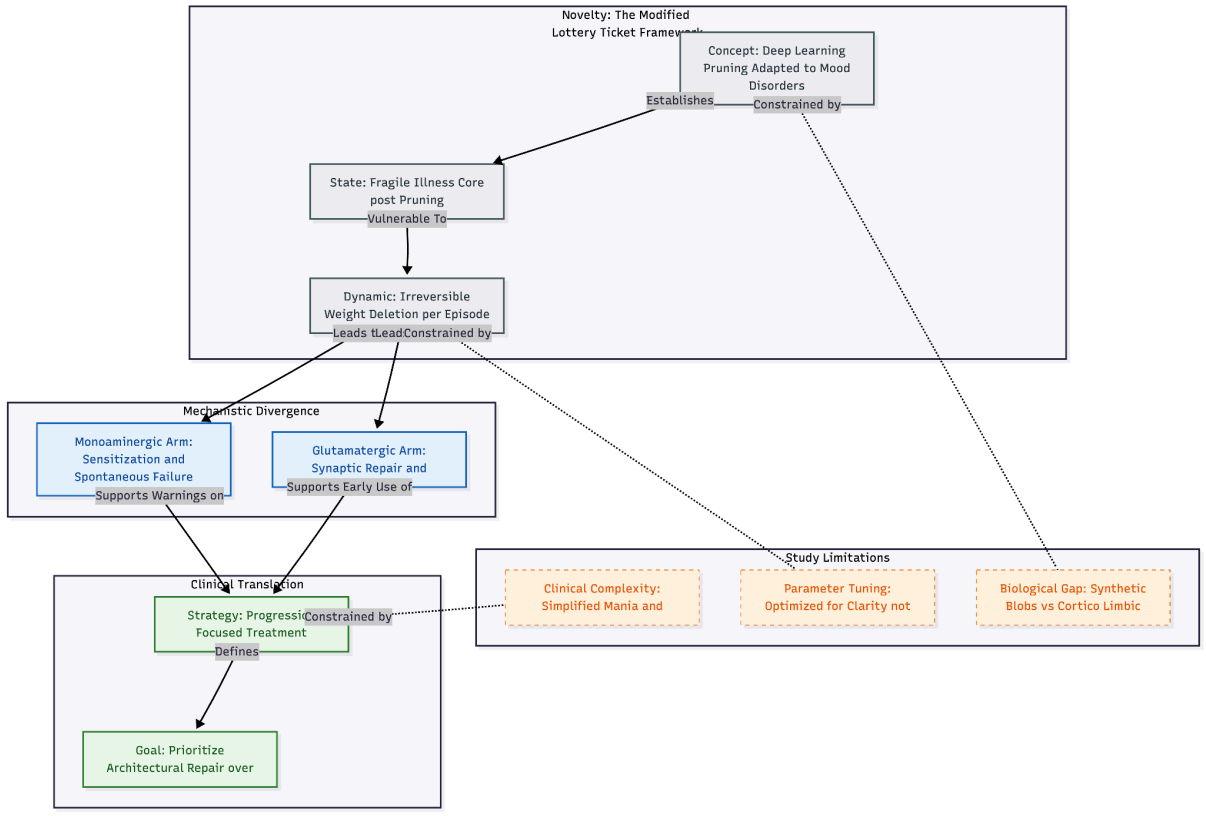

Figure 2: *Conceptual framework and translational implications. This diagram summarizes the study's contribution to the "lottery-ticket" hypothesis of mood disorders. Top: The novel application of neural pruning creates a simulation environment where episode-dependent scarring permanently alters network topology. Middle: This mechanism produces two distinct trajectories: a sensitization pathway (analogous to ineffective monoaminergic treatment) and a resilience pathway (analogous to glutamatergic plasticity). Bottom: These findings support a progression-focused clinical strategy that prioritizes structural repair to halt neuroprogression. Dashed Box: Interpretation is bounded by the simplified nature of the classifier, parameter tuning for theoretical clarity, and the exclusion of complex biological variables such as endocrine feedback or genetic heterogeneity.*

Concluding remarks

Even within these confines, the network repeatedly reproduced clinical themes—higher switch risk under excitatory gain, state-dependent benefit of neurosteroid inhibition, and plasticity-driven escape from kindling. The convergent patterns lend plausibility to a central proposal: the capacity of a treatment to repair or insulate synapses may govern whether episodes set off malignant neuroprogression. Bridging this computational insight with longitudinal imaging, biomarker studies and pragmatic trials will be an essential next step toward therapies that secure not just remission, but long-term stability.

References

[1] Vieta, E., Berk, M., Schulze, T. G., Carvalho, A. F., Suppes, T., Calabrese, J. R., et al. (2018). Bipolar disorders. Nature Reviews Disease Primers, 4, 18008. https://doi.org/10.1038/nrdp.2018.8

[2] Carvalho, A. F., Firth, J., & Vieta, E. (2020). Bipolar Disorder. The New England journal of medicine, 383(1), 58–66. https://doi.org/10.1056/NEJMra1906193

[3] Gijsman, H. J., Geddes, J. R., Rendell, J. M., et al. (2004). Antidepressants for bipolar depression: A systematic review. American Journal of Psychiatry, 161(9), 1537–1547. https://doi.org/10.1176/appi.ajp.161.9.1537

[4] Tondo, L., Vázquez, G., & Baldessarini, R. J. (2010). Mania associated with antidepressant treatment: comprehensive meta-analytic review. Acta psychiatrica Scandinavica, 121(6), 404–414. https://doi.org/10.1111/j.1600-0447.2009.01514.x

[5] Viktorin, A., Lichtenstein, P., Thase, M. E., et al. (2014). The risk of switch to mania in patients with bipolar disorder during treatment with an antidepressant alone and in combination with a mood stabilizer. The American journal of psychiatry, 171(10), 1067–1073. https://doi.org/10.1176/appi.ajp.2014.13111501

[6] Post, R. M. (1992). Transduction of psychosocial stress into the neurobiology of recurrent affective disorder. American Journal of Psychiatry, 149(8), 999–1010. https://doi.org/10.1176/ajp.149.8.999

[7] Post, R. M. (2007). Kindling and sensitization as models for affective episode recurrence, cyclicity, and tolerance phenomena. Neuroscience & Biobehavioral Reviews, 31(6), 858–873. https://doi.org/10.1016/j.neubiorev.2007.04.003

[8] Post R. M. (2020). How to prevent the malignant progression of bipolar disorder. Revista brasileira de psiquiatria (Sao Paulo, Brazil : 1999), 42(5), 552–557. https://doi.org/10.1590/1516-4446-2020-0874

[9] Bender, R. E., & Alloy, L. B. (2011). Life stress and kindling in bipolar disorder: review of the evidence and integration with emerging biopsychosocial theories. Clinical psychology review, 31(3), 383–398. https://doi.org/10.1016/j.cpr.2011.01.004

[10] Jawad, M. Y., et al. (2021). Ketamine for bipolar depression: A systematic review. International Journal of Neuropsychopharmacology, 24(7), 535–541. https://doi.org/10.1093/ijnp/pyab023

[11] Wilkowska, A., Szałach, Ł., & Cubała, W. J. (2020). Ketamine in bipolar disorder: A review. Neuropsychiatric Disease and Treatment, 16, 2707–2717. https://doi.org/10.2147/NDT.S282208

[12] Gunduz-Bruce, H., Lasser, R., Nandy, I., et al. (2020, September). Open-label, Phase 2 trial of the oral neuroactive steroid GABAA receptor positive allosteric modulator zuranolone in bipolar disorder I and II. In Poster presented at: psych Congress.

[13] Marecki, R., Kałuska, J., Kolanek, A., et al. (2023). Zuranolone - synthetic neurosteroid in treatment of mental disorders: narrative review. Frontiers in psychiatry, 14, 1298359. https://doi.org/10.3389/fpsyt.2023.1298359

[14] Frankle, J., & Carbin, M. (2019). The lottery ticket hypothesis: finding sparse, trainable neural networks. International Conference on Learning Representations. https://doi.org/10.48550/arXiv.1803.03635

[15] Berk, M., Kapczinski, F., Andreazza, A. C., et al. (2011). Pathways underlying neuroprogression in bipolar disorder: focus on inflammation, oxidative stress and neurotrophic factors. Neuroscience & Biobehavioral Reviews, 35(3), 804–817. https://doi.org/10.1016/j.neubiorev.2010.10.001

[16] Post, R. M. (2016). Epigenetic basis of sensitization to stress, affective episodes, and stimulants: Implications for illness progression and prevention. Bipolar Disorders, 18(4), 315–324. https://doi.org/10.1111/bdi.12401

[17] Weiss, R. B., Stange, J. P., Boland, E. M., et al. (2015). Kindling of life stress in bipolar disorder: Comparison of sensitisation and autonomy models. Journal of Abnormal Psychology, 124(1), 4–16. https://doi.org/10.1037/abn0000014

[18] Monroe, S. M., & Harkness, K. L. (2005). Life stress, the "kindling" hypothesis, and the recurrence of depression: Considerations from a life-stress perspective. Psychological Review, 112(2), 417–445. https://doi.org/10.1037/0033-295X.112.2.417

[19] Shapero, B. G., Weiss, R. B., Burke, T. A., et al. (2017). Kindling of life stress in bipolar disorder: Effects of early adversity. Behavior Therapy, 48(3), 322–334. https://doi.org/10.1016/j.beth.2016.12.003