CHEMISTRY, PSYCHOLOGY AND AI

AI-Driven Molecular Docking in Neurological Drug Repurposing: Case Studies of Antidepressants, Metformin, and Ketamine

Sementika

Published November 2026

Abstract

More than 57 million people lived with dementia in 2019, and that number is projected to triple by 2050, yet no current drug can stop or reverse its progression. Developing new neurological drugs from scratch takes more than a decade and fails more often than it succeeds. A faster path may already exist in approved medicines whose mechanisms and safety records are known. Artificial intelligence (AI) is making it possible to find new uses for them by improving the speed, accuracy, and scale of molecular docking. This review draws on published literature to examine how AI-driven docking supports the repurposing of three well-characterized drugs: antidepressants, metformin, and ketamine. We argue that this approach offers various benefits. It saves time by screening many protein targets computationally before laboratory work begins. It builds on existing knowledge of how these drugs work, so researchers can focus on discovering new targets. It lowers costs by narrowing the field to the most promising candidates and reducing laboratory and animal experiments. Further advantages include more accurate machine-learning scoring, access to predicted protein structures through tools such as AlphaFold, detection of off-target effects, and the design of safer analogs. Brief case studies show how these advantages apply in practice. Although experimental validation remains necessary, we conclude that AI-driven docking deserves a central role in drug repurposing for neurological disease.

Keywords: molecular docking; artificial intelligence; drug repurposing; neurological disease; antidepressants; metformin; ketamine

1 Introduction

Neurological and psychiatric disorders are among the leading causes of disability worldwide, and their burden is growing as populations age. Depression alone affects an estimated 280 million people, and dementia affects more than 57 million, a figure projected to nearly triple by 2050. Together with Parkinson's disease, stroke, and other conditions, these disorders cost health systems and families enormous sums each year, with much of the burden falling on unpaid caregivers. Yet for many of them, current treatments are only partly effective. Antidepressants fail to produce remission in roughly one third of patients with depression, and no available drug can stop or reverse the progression of major neurodegenerative diseases.

This reflects the wider difficulty of neurological drug development. Bringing a new drug to market typically takes more than a decade and costs well over a billion US dollars, and failure rates for neurological candidates are among the highest of any therapeutic area. The blood-brain barrier, the complexity of brain disease mechanisms, and the lack of reliable biomarkers all compound the problem.

Drug repurposing, the identification of new therapeutic uses for approved medicines, offers a faster and cheaper path. Because the safety, pharmacokinetics, and manufacturing of these compounds are already established, development timelines and late-stage failure risk are substantially reduced. Three well-characterized drugs illustrate this potential. Antidepressants have been linked to neuroprotective and anti-inflammatory effects beyond mood regulation. Metformin, a first-line antidiabetic, has attracted interest because of the connection between insulin resistance, impaired brain energy metabolism, and neurological disease. Ketamine, an NMDA receptor antagonist best known for rapid antidepressant action, is being explored for its effects on synaptic plasticity and neuroinflammation. For each, the key question is which additional molecular targets they act on in the brain.

Answering that question experimentally across hundreds of candidate proteins is slow and costly. Molecular docking addresses this by predicting how a small molecule binds a protein target and estimating the strength of that interaction, allowing many targets to be screened computationally before laboratory work begins. Conventional docking has limitations, including rigid treatment of receptors, imperfect scoring functions, and dependence on experimentally solved structures. Artificial intelligence (AI) helps overcome these. Machine-learning scoring functions improve the accuracy of affinity prediction, and AlphaFold and related tools provide reliable predicted structures for targets that lack crystal structures. AI models can also flag likely off-target interactions and support the design of safer analogs.

Although these approaches are increasingly applied, the evidence on how AI-driven docking supports repurposing of specific neurological drugs remains scattered across the literature. **This review therefore aims to examine how AI-driven molecular docking advances the repurposing of antidepressants, metformin, and ketamine for neurological disease.** Drawing on published studies, I assess the advantages of this approach in saving time, building on known drug mechanisms, and reducing cost. I illustrate these advantages through brief case studies of each drug, and I discuss the remaining need for experimental validation

2 Molecular docking and AI

2.1 Principles of molecular docking

Molecular docking is a computational method that predicts how a small molecule (the ligand) fits into the binding site of a protein. It has two parts. A search algorithm generates many possible poses of the ligand within the binding site, and a scoring function estimates the binding affinity of each pose so that the most favorable can be ranked. The output is a predicted binding mode and an approximate binding energy, usually reported in kcal/mol, with more negative values indicating stronger predicted binding.

Docking is widely used in structure-based drug design because it is fast with experimental screening. A single docking run takes seconds to minutes, which makes it possible to test one drug against hundreds of protein targets, or a whole library of approved drugs against one target. This is the basis of "reverse docking," in which a known drug is screened against many proteins to find new targets, and it is the approach most relevant to repurposing.

2.2 How AI extends docking

Molecular docking is a method used for predicting how a ligand interacts with its protein target: the binding pose and binding affinity. The classical docking programs include AutoDock Vina, Glide, and GOLD. These docking programs use a search method combined with scoring functions, either physical or knowledge-based. There are several ways in which artificial intelligence extends docking.

2.2.1 Improved scoring functions

Conventional scoring functions make use of simplistic formulas for calculating interactions between proteins and ligands (such as hydrogen bonds, van der Waals, electrostatics, and desolvation) that are quick to compute yet not always efficient in sorting compounds based on affinity. Machine learning techniques, including random forests, CNNs (such as GNINA), and graph neural networks, learn interaction patterns based on experimental data sets comprising thousands of proteins and ligands.

2.2.2 Direct pose prediction with deep learning

Instead of searching exhaustively through conformations, models like **DiffDock** , EquiBind and Tank-Bind predict the binding pose directly. Diffusion-based methods start from random ligand placements and iteratively refine them into plausible poses. This can be faster and sometimes more accurate on blind docking, where the binding site is unknown.

2.2.3 Structure prediction when no structure exists

Docking needs a protein structure. **AlphaFold** (2 and 3), RoseTTAFold, and similar tools predict structures for targets with no experimental data. Co-folding models such as AlphaFold 3 and its open analogs (Boltz, Chai) go further, predicting protein and ligand together as a complex, which blurs the line between structure prediction and docking.

2.2.4 Scaling to ultra-large libraries

Docking billions of molecules is computationally prohibitive. **Active learning** and surrogate models solve this: dock a small subset, train a model to predict docking scores, use it to pre-filter the rest, and dock only the promising candidates. This can cut compute by orders of magnitude while recovering most top hits.

2.2.5 Generative design guided by docking

Generative models (VAEs, diffusion models, reinforcement learning) can design new molecules directly inside a binding pocket (e.g., Pocket2Mol, DiffSBDD), using docking scores as a feedback signal to optimize for binding, drug-likeness, and synthesizability.

3 Benefits of AI-driven docking

3.1 Higher accuracy in finding true binders

A review of ML scoring functions for structure-based virtual screening ((Li et al., 2021)) covered studies from 2015 to 2019. Retrospective studies found ML scoring functions gave higher hit rates and potencies than the classical functions they were compared against. Several were also used prospectively and directly found mid-nanomolar binders with novel chemical structures, without any potency optimization

3.2 Better pose prediction, including blind docking

The DiffDock paper (Corso et al., 2023) framed docking as a generative modeling problem rather than a regression task. It reached a 38% top-1 success rate on PDBBind, versus 23% for the best traditional docking and 20% for the best earlier deep learning method. It also provides confidence estimates with high selective accuracy, so users can tell which predicted poses to trust.

3.3 Works on predicted protein structures

Classical and earlier ML methods struggle when the protein comes from a structure predictor like AlphaFold. In the DiffDock results, previous methods reached at most 10.4% accuracy on computationally folded structures, while DiffDock kept 21.7%. This makes docking usable for targets with no experimental structure.

3.4 Massive speed and cost savings

- **Ultra-large screening:** Luttens et al. (2025) trained a classifier on docking results for 1 million compounds. Applied to a 3.5-billion-compound library, the protocol cut the computational cost of structure-based screening by more than 1,000-fold. Experimental testing found ligands for G protein-coupled receptors, including compounds with multi-target activity.
- In the HASTEN work (Sivula et al., 2023), screening 1.56 billion compounds dropped from about four months to about ten days.

4 Case studies

Drug repurposing implies the search for novel indications for approved compounds, which is interesting for the treatment of neurodegenerative diseases and psychiatric disorders as it allows to accelerate the development process and use compounds with known safety profiles. Nowadays, computational approaches, including molecular docking, network pharmacology, molecular dynamics (MD) simulation and AI-driven ranking become increasingly popular in generating the repurposing hypothesis. The question this section tries to address is the following: what is the value of such predictions on their own.

To investigate it, let us consider three examples of drug repurposing from different disorders and using different computational approaches. First, this will be a comparison of antidepressants for Alzheimer's disease: a study using only computational docking of 24 antidepressants with MARK4 (Ahmad et al., 2024) and a study, where docking of tricyclic antidepressants was followed by in vitro testing of protriptyline (Bansode et al., 2014). Second, we will analyze metformin for Alzheimer's disease comparing a network pharmacology and docking study (Shi et al., 2023) and a study combining docking with in vitro MARK4 inhibition assay (Ashraf et al., 2022). Finally, the examples for psychiatric conditions will include an AI-driven pipeline predicting repurposing of cocaine use disorder, which was later

4.1 Case study 1: antidepressants as candidates for Alzheimer's disease

Rationale. Depression and Alzheimer's disease (AD) often co-occur, and SSRIs have shown some benefit for AD symptoms. This makes antidepressants a natural repurposing pool.

Study A: Ahmad et al. (2024), a purely computational pipeline. The authors docked 24 antidepressants against MARK4, a kinase linked to tau pathology in AD, then ran molecular dynamics (MD)

simulations on six top drugs with donepezil as the control. Docking affinities ranged from −5.5 to −9.0 kcal/mol. MD and MM-GBSA analysis pointed to citalopram and mirtazapine as the most promising. The authors conclude that the compounds need in vitro and in vivo follow-up. The pipeline was ADMET filtering, then docking, then MD, then MM-GBSA. There is no experimental step.

Study B: Bansode et al. (2014), docking followed by experiments. The team first screened 140 FDA-approved nervous system drugs by docking, which flagged tricyclic antidepressants against three AD targets. Five tricyclics (protriptyline, amitriptyline, maprotiline, doxepin, nortriptyline) were tested for acetylcholinesterase inhibition, and protriptyline was the strongest, with an IC50 of about 0.06 mM. Isothermal calorimetry, fluorescence quenching, kinetics and atomic force microscopy then supported its inhibition of all three targets, and MD supported the mechanism. It also inhibited glycation-induced amyloid aggregation.

4.2 Comparative evidence for antidepressant repurposing in Alzheimer's disease

4.2.1 Distinct evidence pathways

Although both studies use molecular docking to identify potential drug–target interactions, they provide different levels of evidence. Ahmad et al. (2024) used an entirely computational workflow, progressing from ADMET screening and docking to molecular dynamics (MD) simulations and MM-GBSA analyses. Their findings therefore constitute a ranked, simulation-supported hypothesis. Consistent with this interpretation, the authors noted that the shortlisted compounds require further in vitro and in vivo evaluation.

In contrast, Bansode et al. (2014) combined docking with experimental validation. After generating docking-based predictions, they assessed candidate compounds using biochemical assays, calorimetry, fluorescence quenching, enzyme kinetics, and atomic force microscopy. MD simulations were used primarily to interpret potential mechanisms of interaction. This combination of independent assay formats provides stronger support for inhibitory activity and reduces the likelihood that the findings reflect an artifact of a single experimental method.

4.2.2 Different targets and therapeutic hypotheses

The two studies should not be interpreted as replications because they address different molecular targets and therapeutic hypotheses. Ahmad et al. focused on MARK4, a target implicated in tau pathology. MARK4 phosphorylates tau and may contribute to tau accumulation and tauopathy; phosphorylated MARK4 has also been reported to co-localize with phospho-tau Ser262 in Alzheimer's disease samples.

Bansode et al., however, investigated a multi-target strategy involving acetylcholinesterase (AChE), BACE-1, and amyloid aggregation. Their rationale was that multi-target-directed ligands may be appropriate for a complex disorder such as Alzheimer's disease. The candidate pools also differed substantially: Ahmad et al. screened 24 antidepressants, whereas Bansode et al. evaluated 140 nervous system drugs. Consequently, the apparent convergence of the studies on antidepressants as promising candidates represents limited support rather than direct replication.

4.2.3 Interpretation of docking and molecular dynamics results

Docking scores ranging from −5.5 to −9.0 kcal/mol can rank compounds within a single computational experiment, but they should not be interpreted as measured binding affinities. MD simulations and MM-GBSA analyses provide additional information about whether a predicted binding pose remains

stable over time; however, pose stability does not establish that a compound binds with pharmacologically useful affinity.

Several methodological points require confirmation from the full text of Ahmad et al. These include whether the docking protocol was validated through redocking of a known MARK4 ligand, the duration of the MD simulations, and whether the ADMET analysis considered blood–brain barrier penetration. More broadly, docking-based screening can show poor correspondence between predicted scores and experimental potency and may generate false-positive candidates.

4.2.4 *Potency, exposure, and clinical relevance*

Even the experimentally supported findings require cautious interpretation. Among five tricyclic antidepressants, protriptyline showed the strongest AChE inhibition, with an IC50 of approximately 0.06 mM. By comparison, the metformin–MARK4 study reported an IC50 of 7.05 μM, which is roughly ten-fold lower. Clinically used AChE inhibitors such as donepezil are generally reported to have substantially greater potency; therefore, protriptyline should be evaluated against published donepezil values before its therapeutic potential is judged.

More importantly, in vitro IC50 values do not establish whether comparable concentrations can be achieved in brain tissue at tolerable doses. Evidence of target inhibition in biochemical systems must therefore be complemented by pharmacokinetic, blood–brain barrier, and safety data.

4.3 Case study 2: ketamine for psychiatric disorders

Rationale. Ketamine is an anesthetic that is also used for depression. Computational work on it falls into two distinct strands, so I'd treat them separately.

AI-based repurposing for cocaine use disorder. Gao et al. (2023) addressed a condition with no approved drug. Their pipeline combined a knowledge-graph prediction system (KG-Predict), advisory committee review, retrospective EHR corroboration and genetic and pathway analyses. The EHR step used TriNetX data from over 90 million patients. After propensity matching, cohorts comprised 7,742 CUD patients who received anesthesia and 7,910 CUD patients with depression, with remission within one year as the outcome. Ketamine-exposed patients with depression had a hazard ratio of 4.39 compared with those given other antidepressants or midazolam. The mechanism analysis linked ketamine to CUD-associated genes including BDNF, DRD2, OPRM1 and SLC6A4. The top-10 AI ranking also included drugs that already had clinical trial evidence in CUD. The AI step therefore recovered known leads, and the EHR corroboration carried most of the evidential weight. The coverage of the study also stresses that additional clinical trials are needed.

Docking and network pharmacology in depression. Altê and Rodrigues (2023) compared esketamine and arketamine. They found shared biological targets, particularly NMDA-related ones, along with some poorly explored targets. Yang et al. (2026) examined cognitive impairment in depression. They identified 20 ketamine targets related to depression and cognitive impairment. Five core genes (MAO-A and MAO-B) emerged from the protein interaction network, and docking was used to verify binding. In both studies docking supports a predicted target list and is not independent validation.

Analysis. The two ketamine strands carry different evidential weight and should not be read as converging evidence. In Gao et al. (2023), a knowledge-graph model nominated candidates, experts filtered them, and patient records tested them. The AI ranking was combined with advisory committee review, EHR corroboration and gene and pathway analysis, but the AI step mainly prioritized known leads: the top-10 list also included aripiprazole and quetiapine, which already had

clinical trial evidence in CUD. The evidential weight therefore came from the EHR step, which is large (over 90 million patients) but observational. Propensity matching balances only measured covariates, so confounding by indication remains possible. A coded remission diagnosis is also only a proxy for abstinence, and the hazard ratio of 4.39 is large for an observational comparison, which calls for caution. The investigators themselves emphasized that additional clinical trials are required. The gene-level mechanism analysis is associative, linking ketamine to CUD-related genes through databases, which does not show action at relevant concentrations.

The depression studies (Altê & Rodrigues, 2023; Yang et al., 2026) sit lower on the evidence ladder. They are computational throughout, and docking is a verification step applied to targets already nominated by network analysis, so strong scores to common depression-network nodes (MAO-A, MAO-B, GSK-3β, SIRT1, EGFR) do not separate ketamine's mechanism from that of other antidepressants. A further point is safety: the enantiomer study's preprint reports that both enantiomers show addiction potential, which makes ketamine's use in CUD a risk–benefit question that remission data alone cannot settle. Finally, ketamine is already approved and used for depression, so the case extends an existing drug and says less about AI discovering new uses. Taken together, the ketamine studies show how computational ranking, expert review and real-world data can corroborate a hypothesis, while causality, mechanism, safety in the new population and randomized confirmation remain open.

4.4 Case study 3: metformin as a candidate for Alzheimer's disease

Rationale. Type 2 diabetes (T2D) raises the risk of Alzheimer's disease (AD), and the two conditions share pathways involving insulin signaling, oxidative stress and inflammation. Metformin is an ideal repurposing candidate because its safety profile is well established. In the UKPDS 34 trial, metformin reduced diabetes-related death by 42% and all-cause mortality by 36% in overweight patients with T2D. Its mechanism is thought to involve inhibition of mitochondrial complex I and activation of AMPK, although that mechanism is still debated.

Study A: Shi et al. (2023), a purely computational pipeline. The authors collected metformin targets and AD and T2D disease genes from public databases and identified 115 key targets shared among them. Core targets included TP53, TNF, SIRT1 and CAT. Docking showed the best binding to catalase (CAT), and a Gromacs molecular dynamics simulation showed the CAT–metformin complex converging stably at about 30 ns. The authors propose that metformin acts through oxidative stress, gluconeogenesis and inflammation. The pipeline was network pharmacology, then docking, then MD. There is no experimental step. **Study B: Ashraf et al. (2022), docking followed by experiments.** This study examined microtubule affinity-regulating kinase 4 (MARK4), the same kinase targeted in Case Study 1. Metformin docked to MARK4 with an affinity of −6.9 kcal/mol. An ATPase assay gave an IC50 of 7.05 μM, and a fluorescence binding assay gave a binding constant of 0.6×10^6 M^{-1}. The docking prediction therefore has laboratory support for one proposed target.

Clinical evidence. Luchsinger et al. (2016) ran a pilot trial in amnestic mild cognitive impairment (MCI). It randomized 80 overweight or obese participants without treated diabetes to metformin 1000 mg twice daily or placebo for 12 months, and the authors concluded that a larger trial was warranted. That larger trial, MAP, is a multisite randomized placebo-controlled study of extended-release metformin (2,000 mg a day) in 370 people with amnestic MCI. Its consent materials state that metformin is not approved for preventing cognitive decline and that benefit has not been proven definitively. Observational evidence is inconsistent. One meta-analysis found no significant effect of metformin on the incidence of neurodegenerative diseases overall (OR 1.04), but a higher Parkinson's

disease risk with metformin monotherapy (OR 1.66). A 2026 network meta-analysis of nine cohort studies found no significant difference in Parkinson's risk between antidiabetic drug classes.

Analysis. The two computational studies point to different targets—CAT in one and MARK4 in the other—thereby offering mechanistic hypotheses rather than validating a single target. Their proposed actions are also distinct from metformin's established glucose-lowering mechanisms involving complex I and AMPK. The network-derived CAT finding is supported only by docking and an approximately 30 ns stability simulation, without experimental validation. Ashraf et al. (2022) provide the sole laboratory evidence, and this pertains only to MARK4, reporting an IC50 of 7.05 μM and a binding constant of 0.6×10^6 M^{-1}. The relevance of this potency in the brain remains uncertain: biguanide-mediated complex I inhibition requires concentrations substantially higher than those achieved in blood and tissues at therapeutic doses, and a similar exposure-related uncertainty applies to the micromolar MARK4 effect. candidate; however, its potential benefit in neurodegeneration remains hypothetical until randomized trial results are available.

Clinical research addresses a separate question—whether metformin benefits patients irrespective of the molecular target proposed by docking studies. The pilot study included 80 overweight or obese participants without treated diabetes and concluded that a larger trial was justified. The 370-participant MAP trial continues to examine this question, while its consent materials acknowledge that clinical benefit has not yet been definitively demonstrated. Importantly, both trials involve non-diabetic populations and therefore differ from the T2D association that provides the rationale for metformin's potential neuroprotective role.

Observational evidence also remains inconclusive. One meta-analysis reported no overall association between metformin use and neurodegenerative disease incidence (OR 1.04), although metformin monotherapy was associated with increased Parkinson's disease risk (OR 1.66). Similarly, a 2026 network meta-analysis of nine cohort studies found no significant differences between antidiabetic drug classes. Interpretation of these findings is limited by confounding, particularly because diabetes itself is associated with an approximately 38% higher risk of Parkinson's disease. Overall, metformin's established safety profile and proven efficacy in T2D, including evidence from UKPDS 34, support its plausibility as a repurposing

5 Limitations and challenges

5.1 Limitations of molecular AI docking

The case studies above show that AI-driven docking can generate credible repurposing hypotheses quickly and at low cost, but they also show how far those hypotheses remain from clinical evidence. Only two of the studies reviewed tested their docking predictions in the laboratory, and none established clinical benefit in randomized trials. This section first outlines the limitations of the method itself and then the challenges that must be addressed before computational predictions can reliably support neurological drug repurposing.

1. Predicted binding can form a mechanistic hypothesis but should not be taken for evidence of clinical effectiveness of the drug or its binding to the target. For instance, docking scores are not measurements of binding affinities. While docking scores in the −5.5 to −9.0 kcal/mol range given by Ahmad et al. (2024) could assist in ranking molecules within one experiment, their correlation with experimental potency is weak, and false positives are common. Only when Ashraf et al. (2022) confirmed the activity experimentally using IC50 measure, did −6.9 kcal/mol score for METFORMIN

for MARK4 become useful information. In a similar way, short molecular dynamic simulations can indicate whether a docking pose will remain stable under simulated conditions but cannot confirm binding to the target, let alone the biological significance of interaction. Thus, the CAT-metformin complex simulated for approximately 30 ns by Shi et al. (2023) demonstrates this problem: the stability of a docking pose under the modeled conditions does not mean pharmacological affinity.

2. These approaches are also incomplete representations of the biological environment in which the drugs work. Docking does not take into account the expression of the target, competition with endogenous ligands, post-translational modification, protein-protein interactions, subcellular localization, or changes in cell environment associated with a certain disease.

3. Pharmacokinetic factors are important too but are often insufficiently considered. A drug can demonstrate micromolar activity in an assay, and still not reach necessary concentrations in brain tissues. These problems exists to IC50 of protriptyline of approximately 0.06 mM and to IC50 of metformin for MARK4 of 7.05 μM The blood-brain barrier penetrability, tissue distribution, metabolism, and excretion, thus, should be checked directly, since ADMET screen cannot cover all these issues.

4. The mechanistic interpretation is also limited by off-target activity and the non-specificity of networks derived target lists. Docking can indicate a plausible target while ignoring other interactions important for efficacy and safety of the molecule. In the case of ketamine, addiction potential of its enantiomers cannot be assessed using a docking-derived target list. The same concerns can be raised in ketamine treatment of depression studies: the high docking scores for common nodes of the network like MAO-A, MAO-B, GSK-3β, SIRT1, and EGFR cannot separate ketamine from other antidepressants in terms of mechanisms. In this sense, docking confirms the plausibility of a preselected target list.

5.2 Challenges of AI docking

Beyond the technical limitations of the method, several broader challenges determine whether docking-based predictions can be translated into treatments for neurological disease.

1. Most of the evidence reviewed here stops at computation. Four of the six studies (Ahmad et al., 2024; Shi et al., 2023; Altê and Rodrigues, 2023; Yang et al., 2026) rely on docking, network analysis, and simulation alone, and their authors call for in vitro and in vivo follow-up. Only Bansode et al. (2014) and Ashraf et al. (2022) tested predictions experimentally. Closing this gap is the central challenge for the field, because computational rankings cannot be treated as evidence of activity until they are confirmed

2. A neurological drug must cross the blood-brain barrier at free concentrations comparable to those that were active in vitro. This is uncertain for protriptyline (IC50 ≈ 0.06 mM) and for the micromolar effect of metformin on MARK4 (7.05 μM), yet docking pipelines rarely model brain penetration, efflux transport, or tissue distribution. Repurposing also does not automatically transfer a drug's safety record to a new disease. The addiction potential reported for ketamine enantiomers makes its use in cocaine use disorder a risk-benefit question, and the possible association between metformin and Parkinson's disease remains unresolved.

3. Strong repurposing hypotheses combine docking, network pharmacology, molecular dynamics, genetics, electronic health record data, and clinical trials, as in Gao et al. (2023). Each source has different strengths. In that study, the AI step mainly prioritized candidates that already had clinical trial support, while the observational data carried most of the evidential weight. Weighting and combining such sources rigorously, and reporting them in a standardized way that allows

comparison between studies, remains an open problem. Large EHR datasets allow rapid hypothesis testing, but they remain vulnerable to confounding by indication, and coded outcomes such as remission are only proxies for clinical response. Observational studies of metformin are inconsistent, and the clinical trials (Luchsinger et al., 2016; MAP) enroll non-diabetic participants, who differ from the type 2 diabetes population that motivated the hypothesis.

Even validated target binding does not establish therapeutic benefit. Docking-based mechanisms must ultimately be supported by biochemical assays, cellular and animal studies, pharmacokinetic evidence, and randomized clinical trials.

6 Future directions

Over the next few years, the value of AI-driven docking in neurological repurposing will depend less on gains in speed or scale, which are already substantial, than on how well its predictions are tested and translated .Progress in this field will depend on closing the gap between computational prediction and clinical evidence. The following priorities describe what that would require.

1. Computational hits should be paired with inexpensive orthogonal assays, such as enzyme inhibition, binding measurements, and cell-based tests, before they are reported as promising. The contrast between Ahmad et al. (2024) and Bansode et al. (2014) shows how much stronger the evidence becomes when docking is followed by experiments. Prospective studies that test predictions made in advance, rather than verifying targets already nominated by network analysis, would be especially informative.
2. Docking scores should be combined with predictions of blood-brain barrier permeability, efflux transport, free brain concentration, and metabolism, for example through physiologically based pharmacokinetic modeling, so that candidates are ranked by whether the predicted binding is achievable at tolerable doses. Model development should also focus on target classes that are poorly represented in training data, such as membrane proteins, ion channels, and flexible proteins like tau, using co-folding models and flexible-receptor docking. Reverse docking against proteome-scale structure sets, including predicted structures, could improve off-target and safety prediction, and generative design guided by docking could be used to build safer analogs of drugs with unwanted effects, such as the addiction potential reported for ketamine enantiomers.
3. Combining AI predictions with electronic health record analyses, genetics, and biomarker data, as in Gao et al. (2023), can help prioritize candidates and identify suitable patient groups. Stronger causal methods, such as target trial emulation, and biomarker-guided or adaptive randomized trials are then needed to test the most promising hypotheses, including the ongoing MAP trial of metformin. Funding and regulatory pathways for repurposing off-patent drugs will also matter, since these drugs offer little commercial incentive for late-stage trials.

7 Conclusion

Neurological and psychiatric disorders remain among the leading causes of disability worldwide, while the slow, costly, and high-risk nature of conventional drug development leaves many patients without effective treatments. Drug repurposing offers a more efficient alternative because approved medicines have established safety profiles, pharmacokinetic properties, and manufacturing processes. This review examined the role of AI-driven molecular docking in supporting this approach, using antidepressants, metformin, and ketamine as case studies.

AI-driven molecular docking provides a rapid, cost-effective approach for generating drug-repurposing hypotheses for neurological and psychiatric disorders. As illustrated by case studies involving antidepressants, ketamine, and metformin, this approach enables the computational screening of numerous targets while building on drugs with established safety profiles and mechanisms of action. Advances in machine-learning scoring, deep-learning pose prediction, and predicted protein structures have further increased the speed and scalability of screening and extended its application to targets lacking experimentally determined structures.

The evidence reviewed shows clear advantages. Docking allows many protein targets to be screened computationally before laboratory work begins, and it builds on existing knowledge of how these drugs act, so researchers can concentrate on identifying new targets. Artificial intelligence strengthens this approach in several ways: machine-learning scoring functions improve the identification of true binders, deep-learning models such as DiffDock predict binding poses more accurately, including in blind docking, and AlphaFold-type tools provide structures for targets that lack experimental data.

However, the resulting evidence remains preliminary. Most docking studies reviewed here were limited to computational analyses, with only two evaluating their predictions experimentally. Docking scores and short molecular simulations can rank candidate compounds but do not measure binding affinity and provide limited insight into brain exposure, off-target effects, or safety in a new patient population. Large observational datasets, such as those used in the ketamine study, offer additional support but cannot establish causality, and none of the examined drugs has yet demonstrated clinical benefit in a randomized trial.

The case studies also show how far these hypotheses remain from clinical evidence. Of the docking-based studies reviewed, only two (Bansode et al., 2014; Ashraf et al., 2022) tested their predictions in the laboratory, and even there the reported potencies are in the micromolar to millimolar range, whose relevance to the brain is uncertain. Docking scores are not measured affinities, and short molecular dynamics simulations show that a pose is stable in the model rather than that the drug binds the target in a living system. Docking does not represent cellular context, brain exposure, off-target activity, or safety in a new population, and network-derived targets shared with many other antidepressants cannot distinguish one drug's mechanism from another's. In the ketamine study, the AI step mainly recovered leads that already had clinical trial support, and the evidential weight came from observational electronic health record data, which remain open to confounding. For metformin, the clinical and observational evidence is inconsistent, and none of the drugs examined has yet shown benefit in a randomized trial for the new indication.

These limitations do not weaken the case for AI-driven docking; they define its proper role. It is best used as a tool for generating and prioritizing hypotheses, whose predictions gain evidential weight only when supported by experimental binding data, evidence of adequate brain exposure, and ultimately randomized clinical trials.

AI-driven docking therefore deserves a central role in the early stages of drug repurposing for neurological disease, as a means of choosing which candidates to test rather than as proof that they work. Used in this way, it can shorten the path from existing medicines to new treatments for conditions that currently have few effective options.

References

[1] Ahmad SR, Zeyaullah M, AlShahrani AM, Khan MS, Muzammil K, Ahmed F, et al. (2024). Exploring the most promising anti-depressant drug targeting microtubule affinity receptor kinase 4 involved in Alzheimer's disease through molecular docking and molecular dynamics simulation. *PLoS One*, 19(4), e0301179.

[2] Altê GA, Rodrigues ALS (2023). Exploring the molecular targets for the antidepressant and antisuicidal effects of ketamine enantiomers by using network pharmacology and molecular docking. *Pharmaceuticals*, 16(7), 1013.

[3] Ashraf GM, DasGupta D, Alam MZ, Baeesa SS, Alghamdi BS, Anwar F, et al. (2022). Inhibition of microtubule affinity regulating kinase 4 by metformin: exploring the neuroprotective potential of antidiabetic drug through spectroscopic and computational approaches. *Molecules*, 27(14), 4652.

[4] Bansode SB, Jana AK, Batkulwar KB, Warkad SD, Joshi RS, Sengupta N, et al. (2014). Molecular investigations of protriptyline as a multi-target directed ligand in Alzheimer's disease. *PLoS One*, 9(8), e105196.

[5] Corso G, Stärk H, Jing B, Barzilay R, Jaakkola T (2023). DiffDock: diffusion steps, twists, and turns for molecular docking. In: *Proceedings of the International Conference on Learning Representations (ICLR)*, Kigali, Rwanda. https://doi.org/10.48550/arXiv.2210.01776

[6] Gao Z, Winhusen TJ, Gorenflo M, Ghitza UE, Davis PB, Kaelber DC, et al. (2023). Repurposing ketamine to treat cocaine use disorder: integration of artificial intelligence-based prediction, expert evaluation, clinical corroboration and mechanism of action analyses. *Addiction*, 118(7), 1307–1319.

[7] GBD 2019 Dementia Forecasting Collaborators (2022). Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. *Lancet Public Health*, 7(2), e105–e125.

[8] Li H, Sze KH, Lu G, Ballester PJ (2021). Machine-learning scoring functions for structure-based virtual screening. *WIREs Computational Molecular Science*, 11(1), e1478.

[9] Luchsinger JA, Perez T, Chang H, Mehta P, Steffener J, Pradabhan G, et al. (2016). Metformin in amnestic mild cognitive impairment: results of a pilot randomized placebo controlled clinical trial. *Journal of Alzheimer's Disease*, 51(2), 501–514.

[10] Luttens A, Cabeza de Vaca I, Sparring L, Brea J, Martínez AL, Kahlous NA, et al. (2025). Rapid traversal of vast chemical space using machine learning-guided docking screens. *Nature Computational Science*, 5, 301–312.

[11] Metformin in Alzheimer's Dementia Prevention (MAP). *ClinicalTrials.gov Identifier NCT04098666*. https://clinicaltrials.gov/study/NCT04098666

[12] Ping F, Jiang N, Li Y (2020). Association between metformin and neurodegenerative diseases of observational studies: systematic review and meta-analysis. *BMJ Open Diabetes Research & Care*, 8(1), e001370.

[13] Qin X, Zhang X, Li P, Wang M, Yan L, Bao Z, et al. (2021). Association between diabetes medications and the risk of Parkinson's disease: a systematic review and meta-analysis. *Frontiers in Neurology*, 12, 678649.

[14] Shi X, Li L, Liu Z, Wang F, Huang H (2023). Exploring the mechanism of metformin action in Alzheimer's disease and type 2 diabetes based on network pharmacology, molecular docking, and molecular dynamic simulation. *Therapeutic Advances in Endocrinology and Metabolism*, 14, 20420188231187493.

[15] Sivula T, Yetukuri L, Kalliokoski T, Käsnänen H, Poso A, Pöhner I (2023). Machine learning-boosted docking enables the efficient structure-based virtual screening of giga-scale enumerated chemical libraries. *Journal of Chemical Information and Modeling*, 63(18), 5773–5783.

[16] UK Prospective Diabetes Study (UKPDS) Group (1998). Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). *Lancet*, 352(9131), 854–865.

[17] World Health Organization. *Depressive disorder (depression)* [Internet]. Geneva: World Health Organization. https://www.who.int/news-room/fact-sheets/detail/depression

[18] Yang L, Zhang Q, Zhang Y, Yao K, Chen X, Tian H, et al. (2026). Network pharmacology reveals the potential mechanism of ketamine in improving cognitive impairment in patients with depression. *Alpha Psychiatry*, 27, 56752.