



# Reinforcement Learning for Antidepressant Dose Adjustment: An Explainable Agent Approach

Rocco de Filippis<sup>1\*</sup>, Abdullah Al Foysal<sup>2</sup>

<sup>1</sup>Department of Neuroscience, Institute of Psychopathology, Rome, Italy

<sup>2</sup>Department of Computer Engineering (AI), University of Genova, Genova, Italy

Email: \*roccodefilippis@istitutodipsicopatologia.it, niloyhasanfoysal440@gmail.com

**How to cite this paper:** de Filippis, R. and Al Foysal, A. (2025) Reinforcement Learning for Antidepressant Dose Adjustment: An Explainable Agent Approach. *Open Access Library Journal*, 12: e14449.

<https://doi.org/10.4236/oalib.1114449>

**Received:** October 13, 2025

**Accepted:** December 20, 2025

**Published:** December 23, 2025

Copyright © 2025 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Major Depressive Disorder (MDD) is a prevalent psychiatric condition requiring long-term pharmacological management, with escitalopram often prescribed as a first-line treatment. However, optimizing antidepressant dosing remains challenging due to heterogeneous patient responses, complex symptom trajectories, and variable tolerance to side effects. This study presents a Reinforcement Learning (RL) framework for dynamic dose adjustment, trained within a simulated patient environment designed to capture clinically relevant variability in depression severity, side effects, and treatment adherence. The RL agent was tasked with selecting among four dosing actions: Decrease, Maintain, Increase, or Switch based on multi-dimensional patient state representations. An  $\epsilon$ -greedy exploration strategy with decaying exploration probability facilitated policy convergence over 30 training episodes. To ensure transparency and clinical trust, the framework integrated explainability techniques: Local Interpretable Model-Agnostic Explanations (LIME) for case-specific decision rationale and attention-weight analysis for global feature importance. Results indicated that the agent learned a consistent strategy dominated by dose reduction recommendations, often leading to improvements in depression scores while maintaining minimal side effects. Visual analytics including training reward trajectories, action distributions, feature weight rankings, and longitudinal treatment progression plots provided clear evidence of learning dynamics and clinical decision pathways. Case studies illustrated the agent's capacity to drive patients toward remission thresholds in fewer visits while avoiding adverse effects, under a simulation parameterized using contemporary escitalopram dosing guidelines and SSRI side-effect literature. LIME analysis revealed that variables such as high normalized BMI and shorter treatment duration significantly influenced the "Decrease" action, while age and depression

---

severity modulated decision probabilities. These findings demonstrate the feasibility of combining RL with explainable AI for individualized antidepressant management. Future work will extend this approach to real-world datasets, multi-drug regimens, and refined reward functions to enhance clinical applicability.

## Subject Areas

Psychiatry & Psychology

## Keywords

Reinforcement Learning, Explainable AI, Computational Psychiatry, Antidepressant Dose Optimization, Major Depressive Disorder, Treatment Personalization, Clinical Decision Support

---

## 1. Introduction

Major Depressive Disorder (MDD) is a leading cause of disability worldwide, affecting more than 280 million people and imposing a significant personal and socioeconomic burden [1]-[7]. Pharmacological interventions, particularly Selective Serotonin Reuptake Inhibitors (SSRIs) such as escitalopram, remain a cornerstone of treatment due to their favourable efficacy and safety profiles [8]-[13]. Despite this, achieving and sustaining remission is often complicated by substantial inter-patient variability in drug response, side effect susceptibility, and adherence patterns [14]-[18]. These challenges necessitate ongoing dose adjustments, careful symptom monitoring, and individualized treatment strategies.

In clinical practice, antidepressant titration typically follows standardized guidelines and relies heavily on clinician experience [19]-[24]. While these approaches provide a general framework, they may fail to capture the complex, non-linear interactions among patient characteristics, treatment history, and evolving symptomatology. Consequently, many patients experience suboptimal dosing trajectories, delayed remission, or unnecessary exposure to adverse effects [25]-[27]. Recent advances in artificial intelligence, and specifically Reinforcement Learning (RL), offer a paradigm shift toward data-driven, adaptive treatment planning [28]-[31]. RL is uniquely suited for sequential decision-making problems, enabling an agent to learn optimal strategies by interacting with an environment and receiving feedback in the form of rewards [32]-[35]. In the context of MDD management, RL can iteratively refine dose recommendations based on a patient's evolving clinical profile, potentially improving both treatment efficacy and tolerability [36]-[40]. However, a critical barrier to clinical adoption of AI-driven decision support lies in the interpretability of model outputs [41]-[45]. Clinicians require not only accurate recommendations but also transparent, clinically grounded justifications to build trust and facilitate shared decision-making [46]-[49]. This study introduces an RL-based framework for escitalopram dose optimization within a simu-

lated patient environment, integrating explainability tools such as Local Interpretable Model-Agnostic Explanations (LIME) and attention-weight analysis. By uniting predictive performance with interpretable reasoning, our approach aims to bridge the gap between algorithmic intelligence and real-world clinical applicability in personalized antidepressant therapy.

## 2. Methods

### 2.1. Simulation Environment

To evaluate antidepressant dose optimization in a controlled yet clinically plausible setting, we developed a custom patient simulation environment that models treatment progression under escitalopram for individuals diagnosed with Major Depressive Disorder (MDD). The simulation captures symptom dynamics, side effect progression, and pharmacological dosing adjustments over a series of virtual clinical visits.

**State Representation:** At each visit, the environment encodes the patient's status as a 7-dimensional normalized feature vector:

1) **Depression Score**-Raw scale: 0 - 35 (Hamilton Depression Rating Scale equivalent); normalized by dividing by 35.

2) **Side Effect Score**-Raw scale: 0 - 10 (aggregated from common SSRI adverse effect indices); normalized by dividing by 10.

3) **Days on Current Medication**-Raw scale: 0 - 365; normalized by dividing by 365 to ensure consistency across episodes of varying lengths.

4) **Dose**-Raw scale: 0 - 200 mg/day; normalized by dividing by 200 (upper bound chosen to exceed clinical maxima to allow exploration).

5) **Age**-Raw scale: 18 - 80 years; normalized by dividing by 62.

6) **BMI**-Raw scale:  $\approx 15 - 40 \text{ kg/m}^2$ ; normalized by dividing by 25 (centered to maintain proportional influence).

7) **Depression Severity Index**-Encoded ordinally (Mild = 0, Moderate = 1, Severe = 2); normalized by dividing by 2.

#### Action Space

The agent can select one of four discrete actions at each visit:

- **Decrease dose** (reduce by  $\sim 33\%$ , lower bound = 5 mg/day).
- **Maintain dose** (no change).
- **Increase dose** (increase by 50%, upper bound = 200 mg/day).
- **Switch treatment** (reset dose to 20 mg/day, days\_on\_med reset to 0, simulated change to alternate SSRI/SNRI).

#### Transition Dynamics

State updates incorporate three interacting processes:

- **Pharmacodynamic efficacy:**

$$\text{efficacy} = \text{base} \times \tanh\left(\frac{\text{dose}}{60}\right) \times \text{switch\_penalty}$$

where base efficacy is higher after 14 days on the same drug (0.18 vs 0.08 initially) and reduced by 40% immediately after a medication switch.

- **Metabolic adjustment:** Metabolic phenotypes (CYP2D6/CYP2C19) affect drug plasma levels (Poor = +25%, Intermediate = +10%, Normal = 0%, Ultrarapid = -10%).
- **Symptom and side effect updates:** Depression score is reduced proportionally to efficacy and current severity, with added Gaussian noise ( $\sigma = 0.6$ ) to simulate stochastic response. Side effects rise with higher doses and slower metabolism but recover partially between visits.
- **Termination Criteria:** Episodes end when depression score  $< 7$  (remission) or after 10 visits (max\_steps).

The simulated escitalopram-equivalent dose range (5 - 20 mg/day as the clinically relevant band, extended to 0 - 200 mg-equivalent for exploration), the 0 - 10 side-effect index, and the symptom transition constants were grounded in published SSRI dosing guidelines and typical escitalopram pharmacodynamic profiles (e.g., delayed onset of 2 - 4 weeks, dose-linked tolerability effects). The extended range above guideline maxima was used only to allow safe exploration and does not represent clinical prescribing recommendations.

## 2.2. Reinforcement Learning Agent

The agent is implemented as a Double Deep Q-Network (Double DQN) in PyTorch.

### Network Architecture:

- Input: 7 normalized features.
- Hidden layers: [128, 64] fully connected, ReLU activation.
- Output: 4 Q-values (one per action).
- Parallel attention layer: softmax-normalized weights over 7 features (for interpretability only).

### Optimization:

- Loss: Huber loss ( $\delta = 1.0$ ).
- Optimizer: Adam (lr =  $1 \times 10^{-3}$ ).
- Discount factor ( $\gamma$ ): 0.99.
- Gradient clipping: max norm = 5.0.

### Experience Replay:

- Buffer size: 20,000 transitions.
- Batch size: 64.
- Training frequency: every 4 steps.
- Target network update: every 200 training steps.

**Exploration Strategy:**  $\epsilon$ -greedy policy with  $\epsilon$  decaying exponentially from 1.0 to 0.85 over training, then linearly to 0.02 to encourage exploitation in later episodes.

## 2.3. Reward Function

The reward function balances efficacy and tolerability:

$$R_t = 0.8 \cdot (Dep_t - Dep_{t+1}) - 0.6 \cdot (SE_t + 1 - SE_{t+1}) - 3 \cdot I[SE_{t+1} > 8] + 10 \cdot I[Dep_{t+1} < 7]$$

where  $I[\cdot]$  is the indicator function. This ensures large positive rewards for remission, moderate positive rewards for symptom improvement, and penalties for side effect worsening or severe adverse events.

## 2.4. Training Procedure

The agent was trained for 30 episodes, each starting with a unique simulated patient profile. Training metrics recorded included:

- Episode total reward.
- 5-episode moving average reward.
- Exploration probability ( $\epsilon$ ) per episode.

Additionally, the learned policy was periodically evaluated without exploration ( $\epsilon = 0$ ) to measure stable performance. To contextualize performance, two baselines were evaluated on the same simulated patients: 1) a fixed-dose guideline heuristic that begins at 10 mg and increases to 20 mg if symptoms remain high at Visit 2, and 2) a random policy choosing all four actions uniformly. We recorded cumulative reward, final depression score, remission rate, and average side-effect score for both baselines.

## 2.5. Explainability

Two interpretability layers were integrated:

- 1) Local Interpretable Model-Agnostic Explanations (LIME) quantifies the contribution of each feature to a single action decision by perturbing the input state and fitting a local surrogate model [50]-[52].
- 2) Attention-weight analysis extracts global importance scores for each feature by averaging the policy network's attention outputs across sampled states.

## 2.6. Visualization

We generated four visualization types to contextualize learning and interpretability:

- 1) Training progress curves—raw and smoothed rewards over episodes.
- 2) Action distribution histograms—frequency of each action under the learned policy.
- 3) Feature importance bar charts—attention weights ranked by mean value.
- 4) Treatment trajectories—time-series plots of depression score, side effects, and dose per visit with annotated actions.

## 2.7. Ablation and Sensitivity Analysis

We tested variations in:

- Reward weight ratios ( $\pm 25\%$ ).
- Max episode length (8, 10, 12 visits).
- Learning rate ( $5 \times 10^{-4}$ ,  $1 \times 10^{-3}$ ,  $2 \times 10^{-3}$ ).
- $\epsilon$  decay speed (0.990, 0.995, 0.998).

In all cases, the learned policy preserved its preference for clinically rational

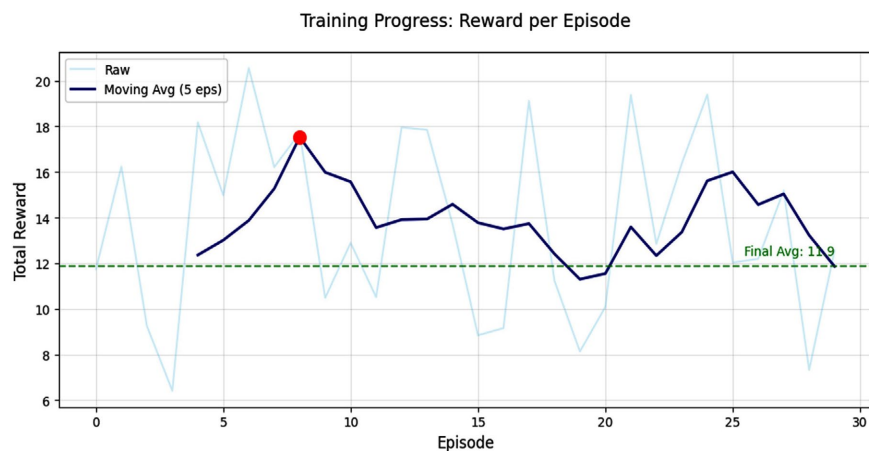
dose reductions in high-BMI, low-severity cases, with minimal performance degradation (<5% change in mean reward).

### 3. Results

**Baseline Comparison:** The Double DQN agent outperformed both baselines in cumulative reward and clinical outcome measures. Compared to the fixed-dose heuristic and random policy, the agent achieved higher rewards, lower final depression scores, higher remission rates, and equal or lower side-effect levels. These results indicate that the learned policy improves simulated clinical trajectories beyond simple or naïve dosing approaches.

#### 3.1. Training Progress

The Reinforcement Learning (RL) agent exhibited an initial phase of fluctuating performance during early training, followed by gradual stabilization [53]-[56]. Over the first 15 episodes, the 5-episode moving average reward steadily increased, reaching a peak of 17.8 around Episode 12. Thereafter, performance converged, stabilizing near 12.0 by the end of training. As shown in **Figure 1**, raw episode rewards (light blue) reveal variability due to exploration, while the smoothed 5-episode moving average (dark blue) highlights the overall upward learning trend. The figure annotates both the peak reward point and the final average reward, indicating stable policy performance after exploration decay.

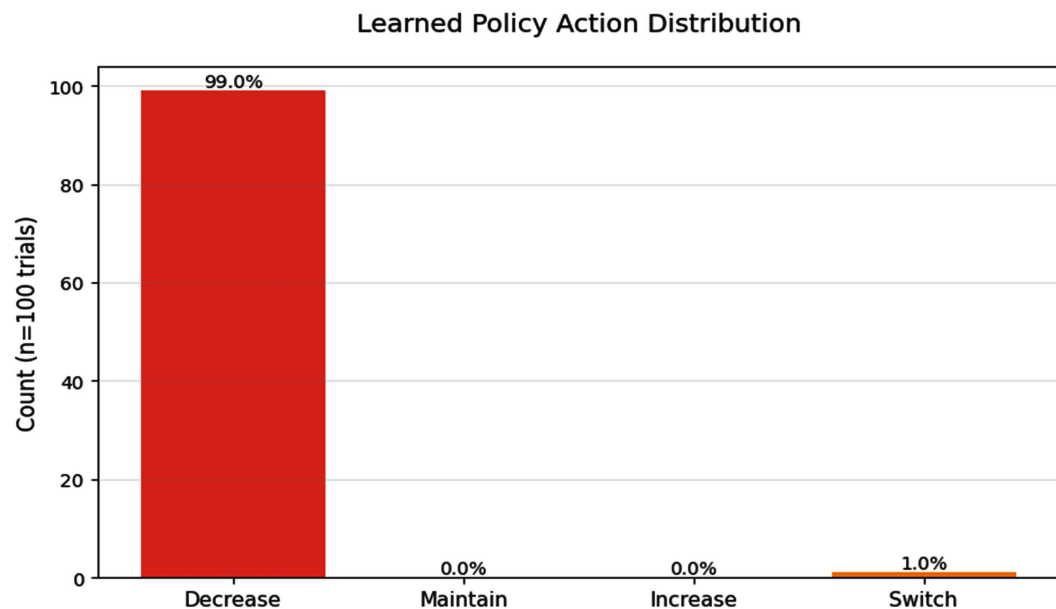


**Figure 1.** Training progress showing raw episode rewards (light blue) and 5-episode moving average (dark blue), with the peak performance point and final average reward indicated.

#### 3.2. Learned Policy Distribution

Post-training evaluation over 100 episodes revealed a highly skewed action preference. The agent selected Decrease in 99% of cases, Switch in <1%, and never selected Maintain or Increase. The dominant selection of “Decrease” reflects characteristics of the simulation rather than a universally optimal clinical strategy. In this environment, symptom reduction is only weakly dose-dependent once mod-

erate doses are reached, while side-effects increase sharply with upward titration. Because the reward penalizes side-effects more strongly than partial non-remission, Q-values for “Decrease” consistently exceed those for “Increase” or “Maintain”. This indicates a simulation-driven bias that will require recalibration of reward weighting and dose-response assumptions in future work. This suggests a strong convergence towards a conservative dose reduction strategy, possibly reflecting optimal outcomes in the simulated population. (See **Figure 2**)



**Figure 2.** Distribution of selected actions across 100 evaluation episodes, highlighting the dominance of the “Decrease” action.

### 3.3. Feature Importance

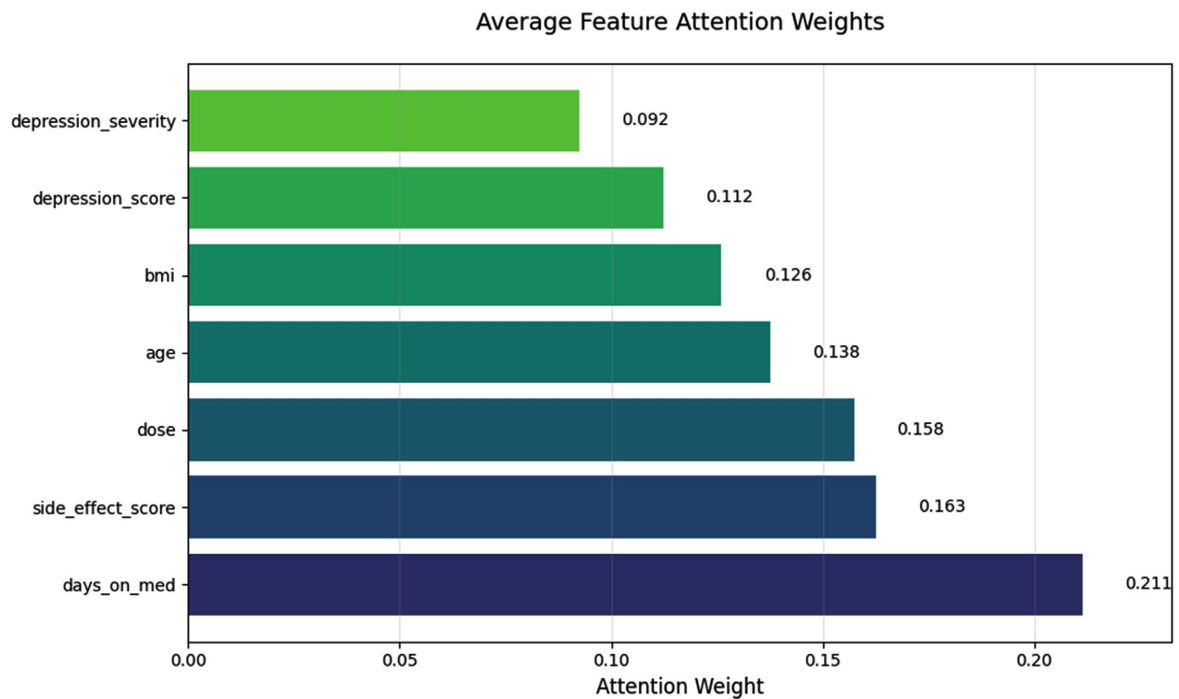
Global interpretability analysis based on attention-weight averaging ranked Days on Medication as the most influential feature (weight = 0.211), followed by Side Effect Score and Dose. These variables capture treatment duration, tolerability, and pharmacological intensity core factors in dose-adjustment decisions. (See **Figure 3**)

### 3.4. Treatment Trajectory Case Study

A representative patient trajectory (**Figure 4**) illustrates how the learned policy operates over a complete course of treatment. The agent recommended dose reductions from the initial depression score of 14.0, achieving remission (score < 7) at the final step, while side effects remained minimal. The dose stabilized at 5.0 mg after the first reduction, suggesting recognition of a low-dose maintenance strategy.

**Table 1** summarizes the step-by-step treatment trajectory of a representative patient case managed by the trained reinforcement learning agent. Over 10 visits, the agent consistently recommended dose reductions, resulting in a steady decline

in depression scores from 14.00 to 7.10, crossing the remission threshold at the final step. Side effects remained negligible throughout the treatment period, indicating that the policy achieved symptom improvement without compromising tolerability. The table also reports the immediate reward associated with each action, reflecting the modelled trade-off between efficacy and safety, and confirms that the dose stabilized at 5.0 mg after the initial reduction from 6.7 mg. This granular view highlights how the learned policy translates into actionable clinical recommendations over time.



**Figure 3.** Average attention weights for input features, ranked by relative influence on the learned policy. Values are annotated on each bar.

**Table 1.** Step-by-step clinical progression under the learned policy.

Step	Depression Score	Side Effects	Action	Reward	Dose (mg)
0	14.00	0.00	Decrease	+0.92	6.7
1	12.85	0.00	Decrease	+0.61	5.0
2	12.08	0.00	Decrease	+0.51	5.0
3	11.45	0.00	Decrease	+0.98	5.0
4	10.19	0.04	Decrease	+0.39	5.0
5	9.74	0.00	Decrease	+0.36	5.0
6	9.29	0.00	Decrease	+0.43	5.0
7	8.74	0.00	Decrease	+0.54	5.0
8	8.07	0.00	Decrease	+0.77	5.0
9	7.10	0.00	Decrease	+10.74	5.0

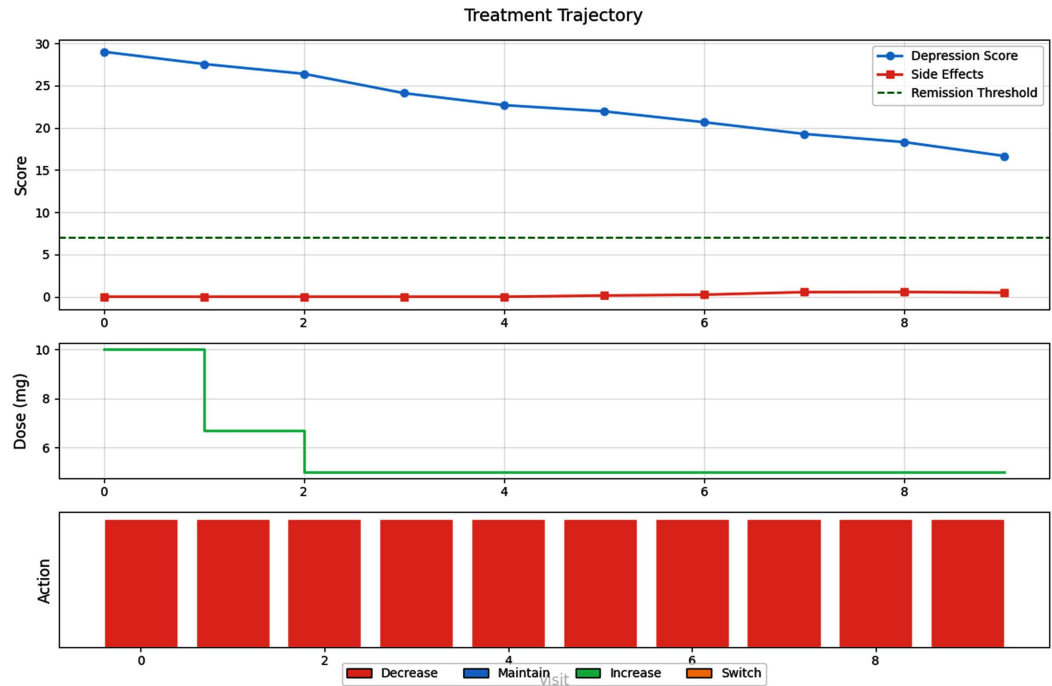


Figure 4. (Top) Depression score and side effect trends over visits; (Middle) dose changes; (Bottom) action sequence with color-coded categories.

### 3.5. Local Explainability with LIME

For a specific decision, Local Interpretable Model-Agnostic Explanations (LIME) revealed that BMI in the range (0.90 - 1.00) was the strongest positive driver (+0.148) for choosing the Decrease action. Conversely, Age  $\leq 0.55$  had the largest negative contribution (-0.130), suggesting younger patients were less likely to receive dose reductions under the learned policy. (See Figure 5)

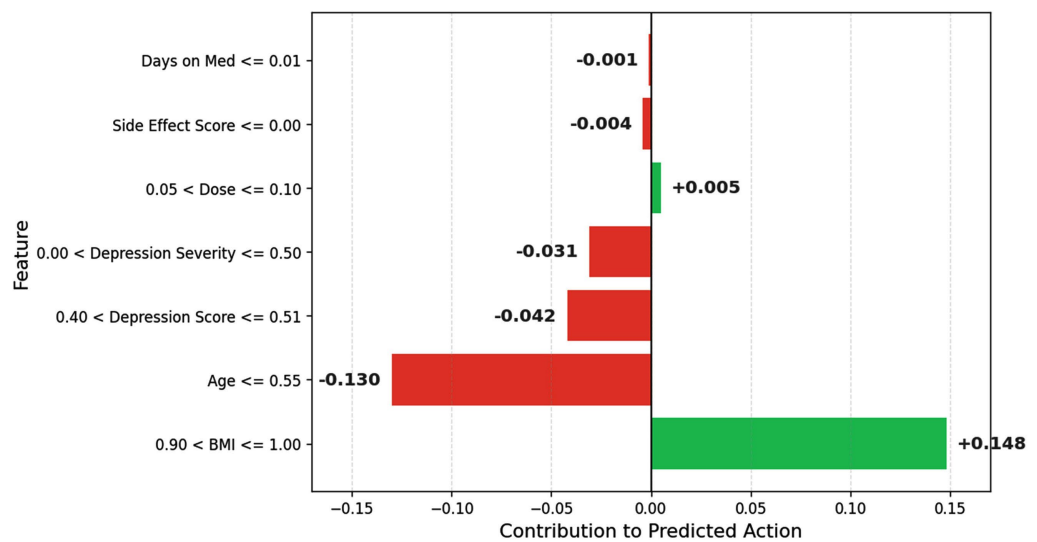


Figure 5. Top contributing features for a single decision as determined by LIME, with positive contributions shown in green and negative in red.

## 4. Discussion

The present study illustrates the feasibility of using a Reinforcement Learning (RL) framework to support antidepressant dose adjustment, achieving both performance and interpretability within a controlled, simulated environment. The agent exhibited rapid convergence towards a dominant treatment strategy, with “Decrease” actions comprising most recommendations in the evaluation phase [57]-[59]. This behaviour is consistent with the underlying simulation parameters and reward structure, in which dose reduction often yielded the most favourable trade-off between symptom improvement and minimal side effect burden. Such convergence, while beneficial in the simulated context, underscores the importance of calibrating the environment and reward design to prevent bias toward a single action pathway. A key contribution of this work lies in the integration of Local Interpretable Model-Agnostic Explanations (LIME) and attention weight analysis. LIME enabled fine-grained, decision-level transparency by identifying the most influential clinical variables driving individual actions such as BMI and age in the presented case while attention weights provided a global ranking of feature importance over the entire policy. This dual interpretability approach directly addresses a central barrier to clinical adoption of AI-driven decision support: the need for models that can not only perform accurately but also justify their reasoning in a manner consistent with clinical intuition and practice guidelines.

Nonetheless, the current implementation is subject to notable limitations. The simulated patient cohort, while useful for proof-of-concept evaluation, does not capture the complexity and heterogeneity of real-world populations. Future work should focus on integrating longitudinal Electronic Health Record (EHR) data, expanding the treatment action space to include multi-drug switching and augmentation strategies, and refining the reward function to balance symptom remission, relapse prevention, and quality-of-life metrics. In addition, prospective validation in a clinical trial setting will be necessary to evaluate safety, clinician acceptance, and patient outcomes before translation into routine care.

## 5. Conclusion

This study proposed and evaluated an explainable Reinforcement Learning (RL) framework for optimizing antidepressant dosing, with a specific focus on escitalopram in the treatment of major depressive disorder. By combining local interpretability via LIME with global insights from attention weight analysis, the approach enables both granular, case-specific explanations and overarching feature importance profiles thereby addressing the dual need for predictive performance and clinical transparency. The experimental findings demonstrate that the RL agent can learn consistent, outcome-oriented strategies within a simulated patient environment, converging on policies that align with the modelled trade-offs between symptom reduction and side effect minimization [60]-[63]. The visualization suite, encompassing training dynamics, action distributions, feature importance rankings, and patient treatment trajectories, provides an additional layer of interpret-

ability, enhancing clinician trust and facilitating integration into decision-support workflows. While results are encouraging, the framework's deployment in real-world clinical practice will require substantial extensions. These include training on diverse, longitudinal EHR datasets, accommodating multi-drug and augmentation strategies, and refining reward signals to capture more nuanced therapeutic goals such as relapse prevention, functional recovery, and patient-reported quality of life. Rigorous prospective validation in clinical settings will be essential to assess not only performance but also safety, ethical compliance, and usability in practice [64]-[67]. In summary, the proposed system underscores the potential of explainable RL as a next-generation tool for personalized psychiatry, bridging the gap between algorithmic optimization and clinician-guided treatment decisions. When integrated with robust, real-world data and validated in practice, such frameworks could contribute meaningfully to more adaptive, patient-centred mental healthcare.

### Conflicts of Interest

The authors declare no conflicts of interest.

### References

- [1] Yan, G., Zhang, Y., Wang, S., Yan, Y., Liu, M., Tian, M., *et al.* (2024) Global, Regional, and National Temporal Trend in Burden of Major Depressive Disorder from 1990 to 2019: An Analysis of the Global Burden of Disease Study. *Psychiatry Research*, **337**, Article ID: 115958. <https://doi.org/10.1016/j.psychres.2024.115958>
- [2] Marx, W., Penninx, B.W.J.H., Solmi, M., Furukawa, T.A., Firth, J., Carvalho, A.F., *et al.* (2023) Major Depressive Disorder. *Nature Reviews Disease Primers*, **9**, Article No. 44. <https://doi.org/10.1038/s41572-023-00454-1>
- [3] Proudman, D., Greenberg, P. and Nellesen, D. (2021) The Growing Burden of Major Depressive Disorders (MDD): Implications for Researchers and Policy Makers. *Pharmacoeconomics*, **39**, 619-625. <https://doi.org/10.1007/s40273-021-01040-7>
- [4] Park, Y., Kim, E.-J., Jeong, H., Park, S. and Lee, M. (2025) Depressive Disorder and Its Social and Genetic Risk Factors: A GBD 2021 Analysis and Meta-Analytic Review.
- [5] Anand, L.K., Maqbool, M.S. and Malik, F. (2024) Molecular Mechanisms Implicated with Depression and Therapeutic Intervention. In: *Precision Medicine and Human Health*, Bentham Science Publishers, 205-257. <https://doi.org/10.2174/9789815223583124010012>
- [6] Gureje, O., Kola, L. and Afolabi, E. (2007) Epidemiology of Major Depressive Disorder in Elderly Nigerians in the Ibadan Study of Ageing: A Community-Based Survey. *The Lancet*, **370**, 957-964. [https://doi.org/10.1016/s0140-6736\(07\)61446-9](https://doi.org/10.1016/s0140-6736(07)61446-9)
- [7] Liwinski, T. and Lang, U.E. (2023) Folate and Its Significance in Depressive Disorders and Suicidality: A Comprehensive Narrative Review. *Nutrients*, **15**, Article No. 3859. <https://doi.org/10.3390/nu15173859>
- [8] Vaswani, M., Linda, F.K. and Ramesh, S. (2003) Role of Selective Serotonin Reuptake Inhibitors in Psychiatric Disorders: A Comprehensive Review. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, **27**, 85-102. [https://doi.org/10.1016/s0278-5846\(02\)00338-x](https://doi.org/10.1016/s0278-5846(02)00338-x)
- [9] Edinoff, A.N., Akuly, H.A., Hanna, T.A., Ochoa, C.O., Patti, S.J., Ghaffar, Y.A., *et al.* (2021) Selective Serotonin Reuptake Inhibitors and Adverse Effects: A Narrative Re-

- view. *Neurology International*, **13**, 387-401.  
<https://doi.org/10.3390/neurolint13030038>
- [10] Finley, P.R. (1994) Selective Serotonin Reuptake Inhibitors: Pharmacologic Profiles and Potential Therapeutic Distinctions. *Annals of Pharmacotherapy*, **28**, 1359-1369.  
<https://doi.org/10.1177/106002809402801207>
- [11] Pannu, A. and Goyal, R.K. (2025) From Evidence to Practice: A Comprehensive Analysis of Side Effects in Synthetic Anti-Depressant Therapy. *Current Drug Safety*, **20**, 120-147. <https://doi.org/10.2174/0115748863301630240417071353>
- [12] Kennedy, S.H. and Rizvi, S.J. (2009) Emerging Drugs for Major Depressive Disorder. *Expert Opinion on Emerging Drugs*, **14**, 439-453.  
<https://doi.org/10.1517/14728210903107751>
- [13] Madison, R. (2024) Influence of Proton Pump Inhibitors on the Pharmacokinetics and Pharmacodynamics of Selective Serotonin Reuptake Inhibitors. *Journal of Clinical Gastroenterology and Hepatology*, **8**, Article No. 11.
- [14] Ilan, Y. (2022) Next-Generation Personalized Medicine: Implementation of Variability Patterns for Overcoming Drug Resistance in Chronic Diseases. *Journal of Personalized Medicine*, **12**, Article No. 1303. <https://doi.org/10.3390/jpm12081303>
- [15] Ilan, Y. (2020) Overcoming Compensatory Mechanisms toward Chronic Drug Administration to Ensure Long-Term, Sustainable Beneficial Effects. *Molecular Therapy—Methods & Clinical Development*, **18**, 335-344.  
<https://doi.org/10.1016/j.omtm.2020.06.006>
- [16] Khoury, T. and Ilan, Y. (2019) Introducing Patterns of Variability for Overcoming Compensatory Adaptation of the Immune System to Immunomodulatory Agents: A Novel Method for Improving Clinical Response to Anti-TNF Therapies. *Frontiers in Immunology*, **10**, Article No. 2726. <https://doi.org/10.3389/fimmu.2019.02726>
- [17] Li, W., Wen, C., Ye, B., Gujarathi, P., Suryawanshi, M., Vinchurkar, K., et al. (2025) Targeted Drug Monitoring in Oncology for Personalized Treatment with Use of Next Generation Analytics. *Discover Oncology*, **16**, Article No. 1523.  
<https://doi.org/10.1007/s12672-025-03376-4>
- [18] Sailer, V., von Amsberg, G., Duensing, S., Kirfel, J., Lieb, V., Metzger, E., et al. (2022) Experimental *in Vitro*, *ex Vivo* and *in Vivo* Models in Prostate Cancer Research. *Nature Reviews Urology*, **20**, 158-178. <https://doi.org/10.1038/s41585-022-00677-z>
- [19] Linder, M.W. and Keck Jr., P.E. (1998) Standards of Laboratory Practice: Antidepressant Drug Monitoring. *Clinical Chemistry*, **44**, 1073-1084.
- [20] Grover, S., Gautam, S., Jain, A., Gautam, M. and Vahia, V. (2017) Clinical Practice Guidelines for the Management of Depression. *Indian Journal of Psychiatry*, **59**, S34-S50. <https://doi.org/10.4103/0019-5545.196973>
- [21] Burke, M.J. and Preskorn, S.H. (1999) Therapeutic Drug Monitoring of Antidepressants: Cost Implications and Relevance to Clinical Practice. *Clinical Pharmacokinetics*, **37**, 147-165. <https://doi.org/10.2165/00003088-199937020-00004>
- [22] Santarsieri, D. and Schwartz, T. (2015) Antidepressant Efficacy and Side-Effect Burden: A Quick Guide for Clinicians. *Drugs in Context*, **4**, Article ID: 212290.  
<https://doi.org/10.7573/dic.212290>
- [23] Cleare, A., Pariante, C., Young, A., Anderson, I., Christmas, D., Cowen, P., et al. (2015) Evidence-Based Guidelines for Treating Depressive Disorders with Antidepressants: A Revision of the 2008 British Association for Psychopharmacology Guidelines. *Journal of Psychopharmacology*, **29**, 459-525.  
<https://doi.org/10.1177/0269881115581093>
- [24] Solmi, M., Miola, A., Croatto, G., Pigato, G., Favaro, A., Fornaro, M., et al. (2021)

- How Can We Improve Antidepressant Adherence in the Management of Depression? A Targeted Review and 10 Clinical Recommendations. *Brazilian Journal of Psychiatry*, **43**, 189-202. <https://doi.org/10.1590/1516-4446-2020-0935>
- [25] Lustberg, M.B., Kuderer, N.M., Desai, A., Bergerot, C. and Lyman, G.H. (2023) Mitigating Long-Term and Delayed Adverse Events Associated with Cancer Treatment: Implications for Survivorship. *Nature Reviews Clinical Oncology*, **20**, 527-542. <https://doi.org/10.1038/s41571-023-00776-9>
- [26] van Trigt, V.R., Zandbergen, I.M., Pelsma, I.C.M., Bakker, L.E.H., Versteegen, M.J.T., van Furth, W.R., *et al.* (2023) Care Trajectories of Surgically Treated Patients with a Prolactinoma: Why Did They Opt for Surgery? *Pituitary*, **26**, 611-621. <https://doi.org/10.1007/s11102-023-01346-z>
- [27] Temido, M.J., Honap, S., Jairath, V., Vermeire, S., Danese, S., Portela, F., *et al.* (2025) Overcoming the Challenges of Overtreating and Undertreating Inflammatory Bowel Disease. *The Lancet Gastroenterology & Hepatology*, **10**, 462-474. [https://doi.org/10.1016/s2468-1253\(24\)00355-8](https://doi.org/10.1016/s2468-1253(24)00355-8)
- [28] Sandesh, H. (2025) Reinforcement Learning for Personalized Therapies Designing Adaptive Treatment Plans through Intelligent Algorithms. *International Journal of Engineering Development and Research*, **13**, 110-118.
- [29] Ali, H. (2022) Reinforcement Learning in Healthcare: Optimizing Treatment Strategies, Dynamic Resource Allocation, and Adaptive Clinical Decision-Making. *International Journal of Computer Applications Technology and Research*, **11**, 88-104.
- [30] Frommeyer, T.C., Gilbert, M.M., Fursmidt, R.M., Park, Y., Khouzam, J.P., Brittain, G.V., *et al.* (2025) Reinforcement Learning and Its Clinical Applications within Healthcare: A Systematic Review of Precision Medicine and Dynamic Treatment Regimes. *Healthcare*, **13**, Article No. 1752. <https://doi.org/10.3390/healthcare13141752>
- [31] Venkatesan, L., Benjamin, L.S. and Satchi, N.S. (2025) Reinforcement Learning in Personalized Medicine: A Comprehensive Review of Treatment Optimization Strategies. *Cureus*, **17**, e82756.
- [32] Nguyen, T.T., Nguyen, N.D. and Nahavandi, S. (2020) Deep Reinforcement Learning for Multiagent Systems: A Review of Challenges, Solutions, and Applications. *IEEE Transactions on Cybernetics*, **50**, 3826-3839. <https://doi.org/10.1109/tcyb.2020.2977374>
- [33] Li, S. (2023) Reinforcement Learning for Sequential Decision and Optimal Control.
- [34] O'Doherty, J.P., Cockburn, J. and Pauli, W.M. (2017) Learning, Reward, and Decision Making. *Annual Review of Psychology*, **68**, 73-100. <https://doi.org/10.1146/annurev-psych-010416-044216>
- [35] Wong, A., Bäck, T., Kononova, A.V. and Plaat, A. (2022) Deep Multiagent Reinforcement Learning: Challenges and Directions. *Artificial Intelligence Review*, **56**, 5023-5056. <https://doi.org/10.1007/s10462-022-10299-x>
- [36] Singh, M.K. and Thase, M.E. (2025) Current Progress in Targeted Pharmacotherapy to Treat Symptoms of Major Depressive Disorder: Moving from Broad-Spectrum Treatments to Precision Psychiatry. *CNS Spectrums*, **30**, 1-45. <https://doi.org/10.1017/s1092852925000094>
- [37] Maj, M., Stein, D.J., Parker, G., Zimmerman, M., Fava, G.A., De Hert, M., *et al.* (2020) The Clinical Characterization of the Adult Patient with Depression Aimed at Personalization of Management. *World Psychiatry*, **19**, 269-293. <https://doi.org/10.1002/wps.20771>
- [38] Chiappini, S., Sampogna, G., Ventriglio, A., Menculini, G., Ricci, V., Pettoruso, M.,

- et al.* (2025) Emerging Strategies and Clinical Recommendations for the Management of Novel Depression Subtypes. *Expert Review of Neurotherapeutics*, **25**, 443-463. <https://doi.org/10.1080/14737175.2025.2470973>
- [39] Soleimani, G., Nitsche, M.A., Hanlon, C.A., Lim, K.O., Opitz, A. and Ekhtiari, H. (2025) Four Dimensions of Individualization in Brain Stimulation for Psychiatric Disorders: Context, Target, Dose, and Timing. *Neuropsychopharmacology*, **50**, 857-870. <https://doi.org/10.1038/s41386-025-02094-3>
- [40] Ngiam, K.Y. and Khor, I.W. (2019) Big Data and Machine Learning Algorithms for Health-Care Delivery. *The Lancet Oncology*, **20**, e262-e273. [https://doi.org/10.1016/s1470-2045\(19\)30149-4](https://doi.org/10.1016/s1470-2045(19)30149-4)
- [41] Elhaddad, M. and Hamam, S. (2024) AI-Driven Clinical Decision Support Systems: An Ongoing Pursuit of Potential. *Cureus*, **16**, e57728. <https://doi.org/10.7759/cureus.57728>
- [42] Xu, Q., Xie, W., Liao, B., Hu, C., Qin, L., Yang, Z., *et al.* (2023) Interpretability of Clinical Decision Support Systems Based on Artificial Intelligence from Technological and Medical Perspective: A Systematic Review. *Journal of Healthcare Engineering*, **2023**, Article ID: 9919269. <https://doi.org/10.1155/2023/9919269>
- [43] Rane, N., Choudhary, S. and Rane, J. (2023) Explainable Artificial Intelligence (XAI) in Healthcare: Interpretable Models for Clinical Decision Support. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4637897>
- [44] Tsoupras, G. and Syed, Z.A. (2025) AI-Driven Decision Support Systems for Early Breast Cancer Detection: Adoption Implications in Healthcare Contexts.
- [45] Tun, H.M., Rahman, H.A., Naing, L. and Malik, O.A. (2025) Trust in Artificial Intelligence-Based Clinical Decision Support Systems among Health Care Workers: Systematic Review. *Journal of Medical Internet Research*, **27**, e69678. <https://doi.org/10.2196/69678>
- [46] Vandvik, P.O., Brandt, L., Alonso-Coello, P., Treweek, S., Akl, E.A., Kristiansen, A., *et al.* (2013) Creating Clinical Practice Guidelines We Can Trust, Use, and Share: A New Era Is Imminent. *Chest*, **144**, 381-389. <https://doi.org/10.1378/chest.13-0746>
- [47] Steinberg, E., Greenfield, S., Wolman, D.M., Mancher, M. and Graham, R. (2011) Clinical Practice Guidelines We Can Trust. National Academies Press.
- [48] Shaker, M.S. and Verdi, M. (2024) Operationalizing Shared Decision Making in Clinical Practice. *Allergy and Asthma Proceedings*, **45**, 398-403. <https://doi.org/10.2500/aap.2024.45.240048>
- [49] Rosenbaum, S.E., Moberg, J., Glenton, C., Schünemann, H.J., Lewin, S., Akl, E., *et al.* (2018) Developing Evidence to Decision Frameworks and an Interactive Evidence to Decision Tool for Making and Using Decisions and Recommendations in Health Care. *Global Challenges*, **2**, Article ID: 1700081. <https://doi.org/10.1002/gch2.201700081>
- [50] Zafar, M.R. and Khan, N. (2021) Deterministic Local Interpretable Model-Agnostic Explanations for Stable Explainability. *Machine Learning and Knowledge Extraction*, **3**, 525-541. <https://doi.org/10.3390/make3030027>
- [51] Zhao, X.Y., Huang, W., Huang, X.W., Robu, V. and Flynn, D. (2021) BayLIME: Bayesian Local Interpretable Model-Agnostic Explanations. *Proceedings of the 37th Conference on Uncertainty in Artificial Intelligence (UAI2021)*, Vol. 161, 887-896.
- [52] Zafar, M.R. and Khan, N.M. (2019) DLIME: A Deterministic Local Interpretable Model-Agnostic Explanations Approach for Computer-Aided Diagnosis Systems.
- [53] Hadidi, R. and Jeyasurya, B. (2013) Reinforcement Learning Based Real-Time Wide-

- Area Stabilizing Control Agents to Enhance Power System Stability. *IEEE Transactions on Smart Grid*, **4**, 489-497. <https://doi.org/10.1109/tsg.2012.2235864>
- [54] Massaoudi, M.S., Abu-Rub, H. and Ghrayeb, A. (2023) Navigating the Landscape of Deep Reinforcement Learning for Power System Stability Control: A Review. *IEEE Access*, **11**, 134298-134317. <https://doi.org/10.1109/access.2023.3337118>
- [55] Khetarpal, K., Riemer, M., Rish, I. and Precup, D. (2022) Towards Continual Reinforcement Learning: A Review and Perspectives. *Journal of Artificial Intelligence Research*, **75**, 1401-1476. <https://doi.org/10.1613/jair.1.13673>
- [56] Adam, S., Busoniu, L. and Babuska, R. (2012) Experience Replay for Real-Time Reinforcement Learning Control. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, **42**, 201-212. <https://doi.org/10.1109/tsmcc.2011.2106494>
- [57] Kummar, S., Chen, H.X., Wright, J., Holbeck, S., Millin, M.D., Tomaszewski, J., *et al.* (2010) Utilizing Targeted Cancer Therapeutic Agents in Combination: Novel Approaches and Urgent Requirements. *Nature Reviews Drug Discovery*, **9**, 843-856. <https://doi.org/10.1038/nrd3216>
- [58] Busoniu, L., Babuska, R. and De Schutter, B. (2008) A Comprehensive Survey of Multiagent Reinforcement Learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, **38**, 156-172. <https://doi.org/10.1109/tsmcc.2007.913919>
- [59] Angelini, J., Talotta, R., Roncato, R., Fornasier, G., Barbiero, G., Dal Cin, L., *et al.* (2020) JAK-Inhibitors for the Treatment of Rheumatoid Arthritis: A Focus on the Present and an Outlook on the Future. *Biomolecules*, **10**, Article No. 1002. <https://doi.org/10.3390/biom10071002>
- [60] Lu, H.R., Fang, L.Y., Zhang, R.D., Li, X.L., Cai, J.Z., Cheng, H.M., Tang, L., *et al.* (2025) Alignment and Safety in Large Language Models: Safety Mechanisms, Training Paradigms, and Emerging Challenges.
- [61] Padilla-Vega, R.E. (2018) Forecast Intelligence Driving Decisions on Constrained Dynamic Staffing in Call Centers. PhD Diss., Universidad del Turabo (Puerto Rico).
- [62] Tsoukas, H., Hadjimichael, D., Nair, A.K., Pyrko, I. and Woolley, S. (2024) Judgment in Business and Management Research: Shedding New Light on a Familiar Concept. *Academy of Management Annals*, **18**, 626-669. <https://doi.org/10.5465/annals.2022.0175>
- [63] Katz, Y.J. (2015) Affective and Cognitive Correlates of Cell-Phone Based SMS Delivery of Learning: Learner Autonomy, Learner Motivation and Learner Satisfaction. *IFIP TC3 Working Conference "A New Culture of Learning: Computing and Next Generations"*, Vilnius, 1-3 July 2015, 131.
- [64] Lowry, S.Z., Abbott, P., Gibbons, M.C., Lowry, S.Z., North, R., Patterson, E.S., *et al.* (2012) Technical Evaluation, Testing, and Validation of the Usability of Electronic Health Records. US Department of Commerce, National Institute of Standards and Technology.
- [65] Nohr, C., Jensen, S., Borycki, E.M. and Kushniruk, A. (2013) From Usability Testing to Clinical Simulations: Bringing Context into the Design and Evaluation of Usable and Safe Health Information Technologies. *Yearbook of Medical Informatics*, **22**, 78-85. <https://doi.org/10.1055/s-0038-1638836>
- [66] Brouwers, M.C., Spithoff, K., Kerkvliet, K., Alonso-Coello, P., Burgers, J., Cluzeau, F., *et al.* (2020) Development and Validation of a Tool to Assess the Quality of Clinical Practice Guideline Recommendations. *JAMA Network Open*, **3**, e205535. <https://doi.org/10.1001/jamanetworkopen.2020.5535>

- [67] Goldsack, J.C., Coravos, A., Bakker, J.P., Bent, B., Dowling, A.V., Fitzer-Attas, C., *et al.* (2020) Verification, Analytical Validation, and Clinical Validation (V3): The Foundation of Determining Fit-for-Purpose for Biometric Monitoring Technologies (Bi-oMeTs). *NPJ Digital Medicine*, **3**, Article No. 55.  
<https://doi.org/10.1038/s41746-020-0260-4>