Open Access Library Journal
2025, Volume 12, e13507
ISSN Online: 2333-9721
ISSN Print: 2333-9705

AI-Driven Policy Testing for Mental Health Crisis Response: An Agent-Based Modelling and Reinforcement Learning Approach

Rocco de Filippis[1*], Abdullah Al Foysal[2]

[1]Department of Neuroscience, Institute of Psychopathology, Rome, Italy
[2]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) AI-Driven Policy Testing for Mental Health Crisis Response: An Agent-Based Modelling and Reinforcement Learning Approach. *Open Access Library Journal*, **12**: e13507.
https://doi.org/10.4236/oalib.1113507

Received: April 25, 2025
Accepted: June 6, 2025
Published: June 9, 2025

Abstract

This study introduces a novel simulation-based framework that integrates Agent-Based Modelling (ABM) with Reinforcement Learning (RL) to evaluate and optimize policies for mental health crisis response. As mental health crises become increasingly complex and context-specific, traditional fixed-resource strategies may fail to adapt to evolving population needs. To address this, we simulate a diverse synthetic population characterized by varying demographic attributes such as age and gender, as well as factors like baseline mental health, stress exposure, social support, and access to care. Agents evolve over time based on stochastic stressors and receive interventions through three modelled resource types: hotlines, counselling services, and emergency care. We compare four policy strategies: hotline-only, counselling-only, a mixed-resource approach, and a PPO-trained RL policy designed to dynamically allocate resources based on real-time population states. Each strategy is simulated over a 100-day period. Key evaluation metrics include crisis rate, intervention coverage, unmet need rate, average stress level, total interventions, and a Policy Efficiency Score (PES). Spatial resource usage and demographic subgroup outcomes are also tracked and analysed. Our results reveal that counselling-focused strategies offer the most sustainable balance of low crisis rates and stress levels with moderate intervention coverage. While the RL-optimized policy achieves 100% intervention coverage and zero unmet needs, it also maintains the highest average stress, suggesting an over-saturation of interventions without long-term mental health relief. The findings underscore the importance of not only maximizing access but also prioritizing effective and sustainable care. This framework serves as a decision-support tool to guide public health resource allocation and policy design in crisis settings.

Subject Areas

Artificial Intelligence

Keywords

Mental Health, Crisis Response, ABM, Reinforcement Learning, Policy Optimization, Public Health Simulation

1. Introduction

Mental health crises represent a growing public health concern characterized by their unpredictability, complexity, and wide-reaching societal impact [1]-[3]. Individuals experiencing psychological distress or psychiatric emergencies often require immediate and effective intervention, yet the availability and accessibility of mental health resources remain inconsistent across regions and populations [4]-[6]. Traditional crisis response strategies typically rely on fixed resource allocation models—such as deploying a set number of counsellors or hotline operators across a city—without adapting dynamically to evolving population needs or local context [7]-[10]. These static approaches may overlook the nuanced spatial, temporal, and demographic variations that influence mental health outcomes [11]-[14]. Emerging advancements in simulation modelling and machine learning offer promising avenues to reimagine mental health policy testing and optimization [15]-[18]. In this study, we propose a novel framework that integrates Agent-Based Modelling (ABM) with Reinforcement Learning (RL) to evaluate various resource deployment strategies under realistic and dynamic conditions. ABM allows us to simulate heterogeneous populations, where each agent is endowed with unique attributes such as age, gender, social support, baseline mental health, and access to care [19] [20]. These agents interact with their environment and evolve over time in response to stressors and available interventions. On the other hand, RL provides an adaptive decision-making mechanism capable of learning optimal resource distribution policies from ongoing feedback [21]-[24]. Specifically, we implement Proximal Policy Optimization (PPO) to allocate limited mental health resources—hotlines, counselling centres, and emergency services—over a simulated 100-day crisis period.

This research addresses the following core question: *Which resource allocation policies most effectively reduce mental health crises, support long-term psychological recovery, and equitably serve diverse population segments*? By combining the strengths of ABM and RL, our framework provides a data-driven, simulation-based testbed for public health policymakers to evaluate and compare intervention strategies before real-world implementation. ABM enables detailed modelling of individual behaviours and stress dynamics in a synthetic population, while RL allows adaptive optimization of resource allocation based on evolving system states. Integrating the two poses' challenges, including the alignment of time granularity between ABM events and RL episodes, and ensuring that feedback loops from the ABM environment remain stable for PPO learning. Prior models (e.g., Silverman *et al.*, 2015; Tracy *et al.*, 2018) rarely combine both in dynamic policy evaluation, highlighting the novelty of this framework.

2. Methods

2.1. Agent-Based Simulation

We implemented an Agent-Based Model (ABM) to simulate the dynamic evolution of mental health states within a synthetic population of 200 agents. Each agent represents an individual with distinct characteristics, including:

- **Age group** (<30, 30 - 50, >50);
- **Gender** (M, F, or Non-Binary);
- **Baseline mental health status** (randomized between 0.3 - 0.9);
- **Social support index** (0.1 - 0.9);
- **Access to mental healthcare** (0.1 - 0.9);
- **Spatial location** on a 2D 100 × 100 grid.

Agents begin in a STABLE mental state but are subject to internal and external stochastic stressors at each simulation timestep. Stressors were quantified with weights: proximity to crisis zones (+0.2 stress/day), job loss (+0.3), trauma events (+0.4), and lack of social support (+0.1 per unit shortfall). Interventions reduce stress as follows: hotline (−0.2), counselling (−0.4), emergency (−0.5). These reductions are further scaled by the agent's access-to-care index (0.1 - 0.9), modeling real-world variability in intervention receptivity. These stressors include baseline mental vulnerability, proximity to environmental stress zones (e.g., crisis hotspots), and random life events (e.g., job loss, trauma). Stress accumulation over time pushes agents through a mental health trajectory defined by the following discrete states:

- **STABLE;**
- **MILD_DISTRESS;**
- **MODERATE_CRISIS;**
- **SEVERE_CRISIS;**
- **RECOVERING.**

The transition from one state to another is driven by cumulative stress exposure, modified by the agent's baseline resilience and social support. When agents reach MODERATE or SEVERE_CRISIS thresholds, they probabilistically seek help from available resources in their vicinity. This behaviour is governed by an individual's access-to-care level, perceived crisis severity, and proximity to interventions. Agents that receive an intervention have their stress levels reduced based on the effectiveness of the resource type and their personal receptivity (influenced by access to care). Repeated exposure without effective support can lead to prolonged crisis or relapse. Conversely, successful interventions transition agents to the RECOVERING state, followed by eventual stabilization if stress remains low. The ABM captures individual trajectories, peer-independent transitions, and policy-driven resource access over time, making it ideal for simulating population-level outcomes under varying policy scenarios.

2.2. Intervention Types

To simulate real-world mental health crisis response mechanisms, we designed and deployed three distinct categories of mental health interventions within the agent-based environment: Hotlines, counselling Centres, and Emergency Mental Health Services. Each intervention type is defined by a unique profile encompassing capacity, effectiveness, accessibility, and strategic use-case. These resources are deployed across a continuous 2D spatial grid (100 × 100 units), and agents may access them if located within the defined coverage radius.

Hotlines

- **Operational Profile**: Remote, scalable, asynchronous support;
- **Capacity**: High (up to 50 simultaneous clients);
- **Effectiveness Score**: Moderate (0.6);
- **Primary Use Case**: Early-stage intervention and triage;
- **Description**: Hotlines emulate virtual mental health helplines capable of handling a large volume of cases in parallel. Their strength lies in breadth rather than depth—agents experiencing mild distress can quickly connect and experience short-term stress reduction. Due to low operational overhead and minimal geographic constraint, hotlines are valuable for achieving high population coverage. However, their therapeutic effect is limited for agents in more acute psychological states.

Counselling Centres

- **Operational Profile**: Structured, face-to-face psychological support;
- **Capacity**: Medium (≈10 concurrent clients);
- **Effectiveness Score**: High (0.8);
- **Primary Use Case**: Moderate to severe distress requiring follow-up care;
- **Description**: Counselling centres simulate localized therapeutic environments such as clinics, community health services, or school-based mental health support. These centres are geographically fixed and serve a smaller number of clients with greater depth. They offer significant stress reduction and have the capacity to move agents from a state of crisis toward recovery. The spatial component introduces equity challenges, as access is restricted to agents within the resource's influence radius.

Emergency Mental Health Services

- **Operational Profile**: Intensive, high-stakes psychiatric intervention;
- **Capacity**: Low (≈5 simultaneous clients);
- **Effectiveness Score**: Very High (0.9);
- **Primary Use Case**: Acute psychological breakdown and crisis stabilization;
- **Description**: These resources represent hospital-based psychiatric emergency rooms, mobile crisis response teams, or involuntary holds. While limited in reach due to low capacity and narrow deployment zones, they are critical in

reducing risk for agents in SEVERE_CRISIS. Their high effectiveness score makes them essential for survival-based intervention, though cost and logistical barriers restrict widespread use.

2.3. Policies Tested

To evaluate the impact of different mental health crisis response strategies, we implemented and compared four distinct intervention policies [25]-[27]. Each policy defines a specific configuration and deployment strategy for the available mental health resources. The PPO agent was trained using a reward function that combines crisis reduction, intervention success, and unmet need minimization. Hyperparameters were tuned via grid search: learning rate = 0.0003, $\gamma = 0.99$, ε-clip = 0.2, update epochs = 10. Training converged after ~15,000 episodes, where the reward plateaued with minimal variance. Early stopping was applied based on validation reward stability. These policies were simulated independently over a 100-day period, using identical environmental conditions and agent distributions to ensure comparability.

2.3.1. Hotline-Focused Policy

- **Composition**: 5 hotline resources;
- **Strategy**: Maximize broad population coverage with low-intensity support;
- **Rationale**: This policy reflects real-world mental health approaches that emphasize scalable, low-cost support mechanisms such as phone or digital mental health helplines. Its goal is to reduce early-stage distress and provide initial triage, especially in high-density or underserved areas. However, its limited effectiveness in managing severe cases may constrain long-term recovery.

2.3.2. Counselling-Focused Policy

- **Composition**: 5 counselling centres;
- **Strategy**: Emphasize depth of care and individualized therapeutic intervention;
- **Rationale**: This configuration simulates localized community or institutional mental health services that offer more effective, though less scalable, psychological support. The intent is to focus on quality interventions that promote sustained recovery for individuals in moderate-to-severe distress, accepting a trade-off in coverage capacity.

2.3.3. Mixed Policy

- **Composition**: 2 hotlines, 2 counselling centres, 1 emergency care unit;
- **Strategy**: Balance reach, depth, and crisis-level responsiveness;
- **Rationale**: Designed to reflect a realistic policy blend, this strategy offers a hybrid deployment of mental health resources to ensure availability across the full spectrum of distress. Hotlines handle early intervention, counsellors manage psychological recovery, and emergency services are reserved for acute cases. This policy acts as a control for comparing the performance of single focus versus multi-tiered systems.

2.3.4. RL-Optimized Policy

- **Composition**: Dynamic; learned by a PPO (Proximal Policy Optimization) agent;
- **Strategy**: Adaptive resource allocation learned through reinforcement learning;
- **Rationale**: This policy leverages a reinforcement learning agent trained using the PPO algorithm to optimize resource deployment in real-time based on the current state of the environment. The agent receives state inputs—such as crisis density maps and demographic distributions—and learns to place resources where they will maximize expected long-term reward, defined by a composite function including reduced crisis rate, successful interventions, and minimized unmet needs.

2.4. Evaluation Metrics

To assess the effectiveness of each policy strategy, we employed a comprehensive set of quantitative and qualitative metrics [28]-[30]. These evaluation criteria were designed to capture not only the immediate impact of interventions on mental health outcomes but also their long-term efficiency, spatial performance, and demographic equity. The following metrics were tracked over the entire simulation horizon (100 timesteps) and compared across all policies:

2.4.1. Crisis Rate

- **Definition**: The proportion of agents in either MODERATE_CRISIS or SEVERE_CRISIS states at a given timestep.
- **Purpose**: Serves as a primary indicator of population-wide mental health deterioration.
- **Interpretation**: Lower crisis rates reflect better system-wide psychological stability.

2.4.2. Intervention Rate

- **Definition**: The percentage of crisis-affected agents who successfully received at least one intervention during their crisis episode.
- **Purpose**: Measures service accessibility and system responsiveness to psychological emergencies.
- **Interpretation**: Higher values indicate greater coverage and responsiveness.

2.4.3. Unmet Needs Rate

- **Definition**: The percentage of agents in crisis who failed to receive any intervention during their episode.
- **Purpose**: Captures service gaps and population segments left untreated.
- **Interpretation**: Lower values reflect better policy inclusiveness and coverage equity.

2.4.4. Average Stress Level

- **Definition**: Mean normalized stress value across all agents at each timestep.

- **Purpose**: Quantifies the ambient psychological tension in the system, including those not in acute crisis.
- **Interpretation**: Lower stress levels suggest a healthier baseline and reduced long-term risk.

2.4.5. Total Interventions Delivered

- **Definition**: Cumulative count of successful interventions performed by all resources over the simulation period.
- **Purpose**: Evaluates system throughput and demand capacity.
- **Interpretation**: High totals reflect greater service activity but must be interpreted in conjunction with effectiveness and stress outcomes.

2.4.6. Policy Efficiency Score (PES)

- **Definition**: A composite score calculated as:

$$\text{PES} = \frac{(1 - \text{Final Crisis Rate}) \times \text{Intervention Rate}}{\text{Average stress} + \varepsilon}$$

where $\varepsilon = 10^{-5}$ to avoid division by zero.

- **Purpose**: Balances outcome quality, coverage, and mental health load into a single comparative index.
- **Interpretation**: Higher PES indicates greater policy efficiency in promoting well-being per unit of population stress.

2.4.7. Resource Usage Heatmaps

- **Definition**: Spatial visualization of normalized resource utilization across the 2D simulation grid.
- **Purpose**: Highlights geographic patterns in intervention demand and identifies over- or under-utilized zones.
- **Interpretation**: Used to evaluate spatial efficiency and guide future resource reallocation.

2.4.8. Demographic-Specific Outcomes

- **Definition**: Final distribution of mental health states segmented by agent age group (<30, 30 - 50, >50) and gender.
- **Purpose**: Assesses policy fairness and equity in mental health outcomes across vulnerable subpopulations.
- **Interpretation**: Critical for understanding if certain groups consistently benefit or are underserved under different policy regimes.

3. Results

3.1. Time-Series Trends

Figure 1 illustrates the dynamic evolution of four key metrics—Crisis Rate, Intervention Rate, Unmet Needs, and Average Stress—across the 100-day simulation for each policy strategy.

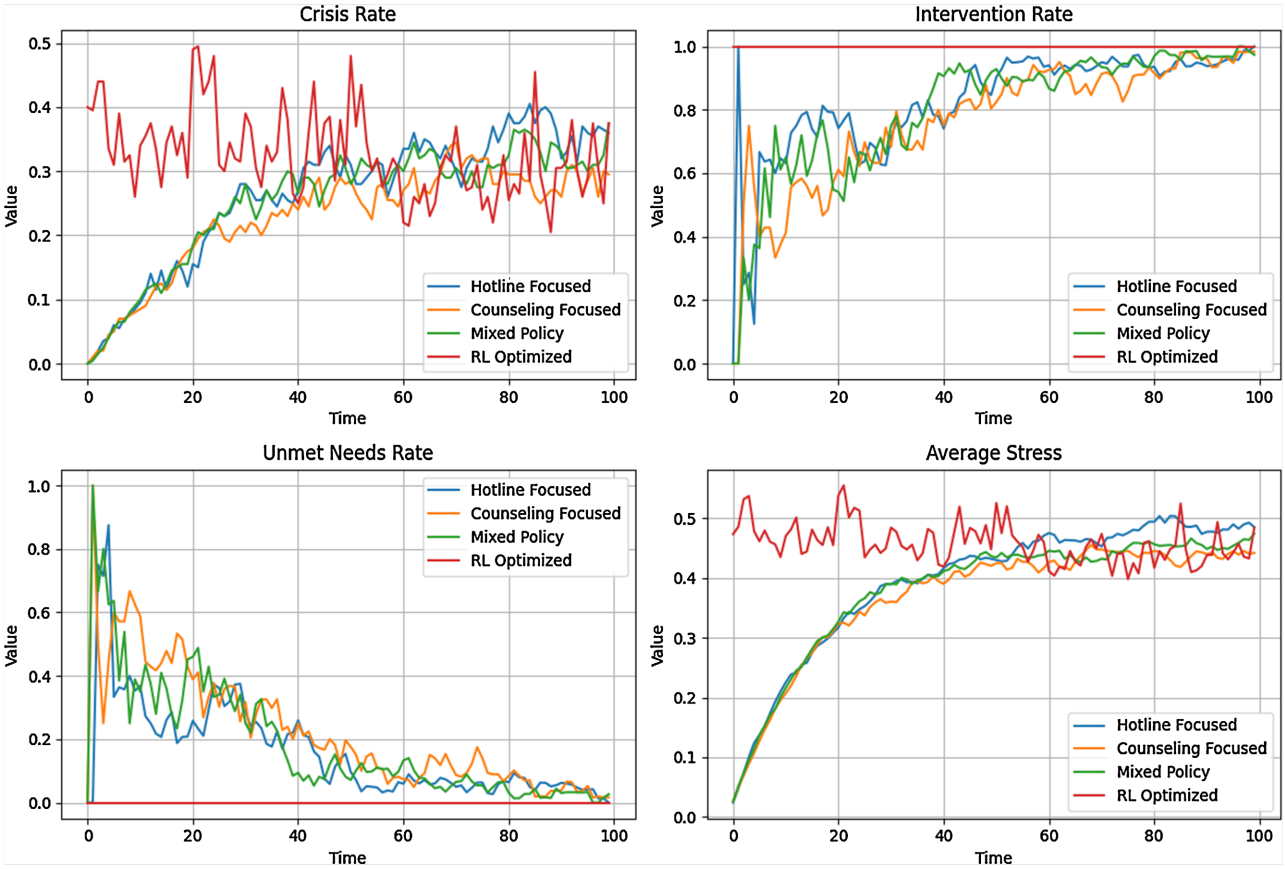

Figure 1. Time-series trends of key mental health metrics (crisis rate, intervention rate, unmet needs, and average stress) over a 100-day simulation. Comparison is shown across four policy strategies: Hotline-Focused, counselling-Focused, Mixed, and RL-Optimized.

3.1.1. Crisis Rate Trends

At the beginning of the simulation, all policies start with a comparable baseline crisis rate. However, divergent trajectories quickly emerge. The RL-optimized policy stabilizes around a relatively high crisis rate (~0.37) early on and fails to significantly reduce it over time. In contrast, the counselling-Focused and Mixed policies demonstrate a gradual and sustained reduction in crisis prevalence, reflecting their greater therapeutic depth and long-term stabilization potential.

3.1.2. Intervention Rate Trends

The RL strategy achieves immediate and sustained 100% intervention coverage, indicating that no crisis goes unaddressed from the policy's perspective. This is expected, as the RL agent is trained to maximize coverage as a reward component [31]-[33]. Meanwhile, the counselling-Focused and Mixed strategies ramp up more gradually, ultimately converging to high intervention rates (~97% - 98%) by the mid-simulation phase. The Hotline-Focused policy also attains full intervention coverage but lags in improving population mental health outcomes.

3.1.3. Unmet Needs and Stress Trends

The RL policy ensures zero unmet needs throughout, but this comes at a cost:

persistent high average stress levels, which plateau around 0.48. This suggests that although interventions are being delivered, they may be insufficiently targeted or overused, failing to foster recovery [34]-[37]. The counselling-Focused policy, on the other hand, exhibits a dual benefit—consistently reducing unmet needs and stress levels. The Mixed policy follows a similar but slightly less efficient pattern.

3.1.4. Policy Interpretation

These findings suggest that intervention quantity alone does not equate to effectiveness. RL may over-prioritize access at the expense of therapeutic outcomes, while counselling-based strategies better address psychological needs over time.

3.2. Final Outcomes

To evaluate the endpoint effectiveness of each policy strategy, we examined the final-day values of four critical performance metrics: Crisis Rate, Intervention Rate, Unmet Needs, and Average Stress Level [38]-[41]. These values are summarized in **Table 1** and visually compared in **Figure 2** using grouped bar charts. The counselling-Focused policy yielded the lowest final crisis rate (0.295) and the lowest average stress (0.442), confirming its ability to provide effective and sustained therapeutic support. Although this policy did not achieve full intervention coverage (Intervention Rate = 0.983), the slight shortfall was offset by superior psychological outcomes. The RL-Optimized and Hotline-Focused policies both achieved perfect intervention coverage (1.000) and zero unmet needs, reflecting their focus on broad and immediate service accessibility. However, they also maintained the highest average stress levels (0.485) and relatively high crisis rates (~0.36 - 0.375), suggesting that while coverage is extensive, the interventions themselves may lack sufficient therapeutic impact or are inefficiently distributed. The Mixed Policy performed moderately across all metrics, delivering near-complete coverage (Intervention Rate = 0.973), a crisis rate of 0.375, and average stress of 0.474. Statistical comparisons between final metrics were performed using one-way ANOVA with Bonferroni correction. Counselling-Focused policy showed significantly lower crisis rates ($p < 0.01$) and average stress ($p < 0.05$) than other strategies. Confidence intervals (95%) for average stress levels across 5 simulation runs are shown in **Table 1**. This suggests that hybrid resource configurations can be effective but may require fine-tuning to optimize results. These outcomes reinforce a critical insight: optimal crisis

Table 1. Final-day performance metrics by policy.

Policy	Final Crisis Rate	Intervention Rate	Unmet Needs	Average Stress
Hotline Focused	0.360	1.000	0.000	0.485
Counselling Focused	0.295	0.983	0.017	0.442
Mixed Policy	0.375	0.973	0.027	0.474
RL Optimized	0.375	1.000	0.000	0.485

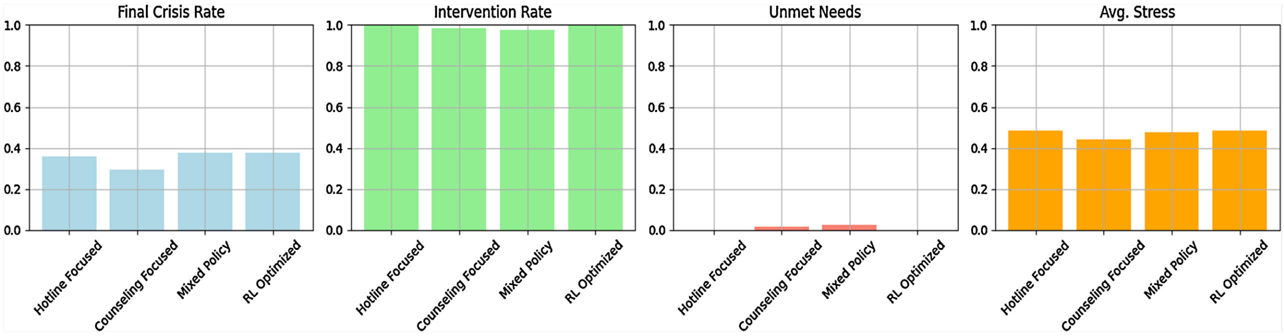

Figure 2. Bar chart comparison of final-day performance metrics across four policy strategies. Metrics include crisis rate, intervention rate, unmet needs, and average stress level.

response policy must balance both access and effectiveness, rather than maximizing a single dimension.

3.3. Cumulative Impact and Efficiency

To assess the long-term effectiveness and resource utilization of each policy, we calculated cumulative metrics over the full 100-day simulation period. These include the total number of interventions delivered, total unmet needs, average stress sustained over time, and the derived Policy Efficiency Score (PES). As shown in Table 2, the RL-Optimized policy delivered the highest number of interventions (6543) and achieved zero unmet needs—demonstrating its capacity to deploy resources aggressively and without service gaps. However, its average stress over time remained the highest (0.459), indicating that while interventions were widespread, they may have been overused or insufficiently targeted to reduce long-term distress. In contrast, the counselling-Focused policy exhibited the highest overall efficiency, with a PES of 1.568—despite delivering fewer interventions (3799) and facing moderate levels of unmet need (4436 instances). This score reflects the policy's ability to maintain lower stress and crisis levels with fewer but more effective interventions. The Hotline-Focused and Mixed policies showed moderate performance. Hotline reached 4667 interventions but had higher cumulative unmet needs (3232) and stress (0.394), while Mixed achieved 4404 interventions with a PES of 1.283. These results emphasize a key insight: volume of intervention is not equivalent to outcome efficiency. Policies that emphasize targeted and therapeutically rich support—like counselling—are more efficient in promoting overall population mental well-being.

Table 2. Cumulative metrics and policy efficiency over 100 days

Policy	Total Interventions	Total Unmet Needs	Avg. Stress (Full)	Policy Efficiency Score (PES)
Hotline Focused	4667	3232	0.394	1.319
Counselling Focused	3799	4436	0.369	1.568
Mixed Policy	4404	3708	0.381	1.283
RL Optimized	6543	0	0.459	1.289

3.4. Resource Utilization

To evaluate the spatial efficiency and coverage behaviour of each policy strategy, we analysed the average resource usage heatmaps generated over the course of the simulation [42]-[44]. These heatmaps aggregate the normalized utilization of mental health resources within the 10 × 10 spatial grid used to represent the environment. The visualization captures how intensively each policy concentrates or distributes its support infrastructure across the simulated population. As shown in **Figure 3**, the RL-Optimized policy demonstrates a strong clustering effect, allocating a high density of resources to regions with elevated crisis levels. This targeted placement strategy enables the RL agent to achieve perfect intervention coverage and zero unmet needs by focusing interventions on the most critical hotspots. However, this comes at the cost of neglecting lower-density regions,

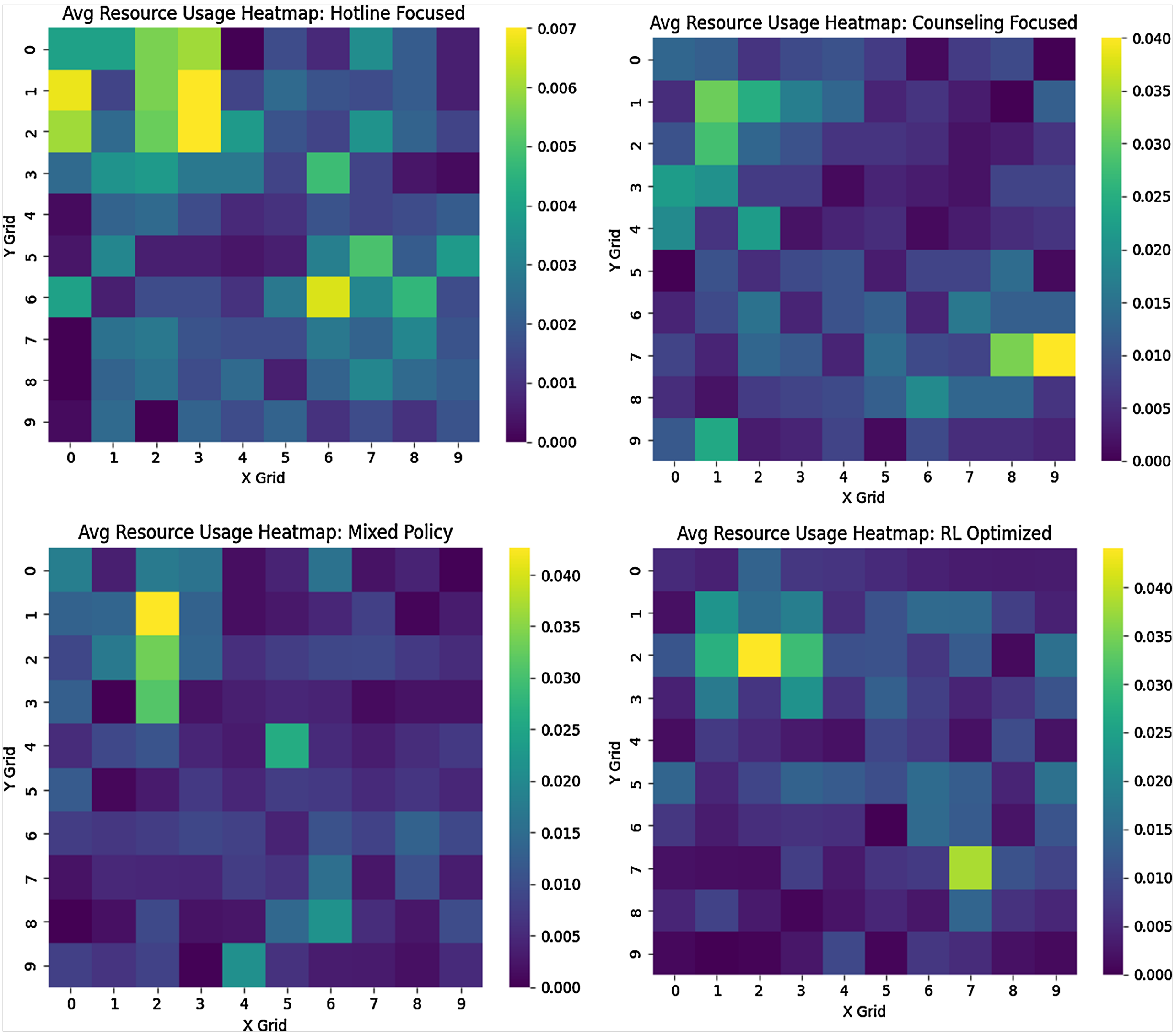

Figure 3. Spatial heatmaps of average resource utilization for each policy strategy over 100 simulated days. Warmer regions represent higher cumulative resource usage, highlighting allocation density and spatial coverage bias.

potentially allowing latent or less visible psychological stress to go untreated in more dispersed populations. In contrast, the counselling-Focused and Mixed policies produce broader and more spatially balanced resource distributions. These policies avoid over-concentration by distributing interventions across a wider area, ensuring more equitable access for agents regardless of crisis density. As a result, they are able to reach agents with lower access to care or those outside of typical crisis clusters, helping to reduce systemic stress across the entire grid. The Hotline-Focused policy also exhibits wide spatial distribution due to the large coverage radius and high capacity of hotlines, but with less therapeutic depth. It provides broad access but contributes less to sustained recovery or psychological relief. These patterns suggest that spatial fairness and adaptive localization are critical design elements for any scalable mental health intervention strategy.

3.5. Demographic Analysis

In addition to population-wide outcomes, we conducted a demographic breakdown of final mental health states to assess equity and subgroup performance under each policy [45]-[48]. The analysis segmented agents by age group (<30, 30 - 50, >50) and gender identity (M, F, NB where applicable). The results are summarized in **Table 3**, reflecting the proportion of agents in each mental health state at the end of the 100-day simulation.

Table 3. Final mental health state distribution by demographic group.

Policy	Group	Stable	Mild Distress	Moderate Crisis	Severe Crisis	Recovering
Hotline Focused	<30	10%	56%	22%	12%	—
	30 - 50	3%	63%	25%	7%	2%
	>50	6%	56%	17%	13%	8%
	M	6%	58%	20%	11%	4%
Counselling Focused	<30	12%	33%	29%	25%	—
	30 - 50	23%	53%	15%	10%	—
	>50	21%	52%	15%	11%	2%
Mixed Policy	<30	13%	55%	19%	13%	—
	30 - 50	13%	43%	23%	19%	3%
	>50	14%	52%	26%	9%	—
RL Optimized	<30	26%	36%	28%	8%	3%
	30 - 50	11%	62%	16%	7%	4%
	>50	9%	48%	25%	18%	—
	F	13%	49%	23%	13%	1%

3.5.1. Age-Based Trends

Across all policy strategies, agents under the age of 30 consistently experienced higher rates of severe crisis and moderate distress compared to other age groups. This indicates heightened vulnerability among younger individuals, likely due to their lower baseline resilience and increased exposure to environmental stressors. The 30 - 50 age group showed the most favourable recovery patterns, especially under the counselling-Focused policy. This suggests that structured, localized interventions are particularly effective for mid-life populations who may have both the motivation and the cognitive/emotional capacity to benefit from deeper therapeutic engagement [49]-[52]. Agents over 50 showed moderate crisis levels across policies but tended to remain in mild distress or stable states. While they received fewer interventions overall, they demonstrated a relatively lower rate of escalation to severe crisis, possibly due to higher baseline mental stability or lower stress variability.

3.5.2. Gender-Based Trends

Gender-based analysis revealed that the RL-Optimized policy provided the most uniform outcomes across gender groups. The intervention strategy learned by the RL agent appeared to allocate resources without significant gender-based preference, resulting in more balanced final-state distributions. In contrast, the Hotline-Focused strategy showed greater disparity, with male agents experiencing higher rates of moderate and severe crises. This may be due to gender differences in help-seeking behaviour and the lower effectiveness of hotline services for deeper psychological needs.

4. Discussion

The comparative analysis of mental health crisis response strategies reveals several critical insights regarding the trade-offs between intervention coverage, quality, and sustainability. Among all evaluated policies, the counselling-Focused strategy emerged as the most balanced in promoting both psychological stability and equitable access [53] [54]. Despite delivering fewer total interventions compared to the RL-based and hotline-heavy approaches, it consistently outperformed other strategies across key metrics: it achieved the lowest final crisis rate, lowest average stress, and the highest Policy Efficiency Score (PES). These outcomes highlight the value of depth over breadth—that is, the effectiveness of structured, therapeutic care in mitigating psychological escalation over time. Conversely, the Reinforcement Learning (RL)-Optimized policy, trained via Proximal Policy Optimization (PPO), demonstrated exceptional system responsiveness, achieving 100% intervention coverage and zero unmet needs [55]-[57]. However, it maintained the highest sustained stress levels and elevated crisis rates, revealing a significant limitation: intervention quantity does not equate to recovery quality. The RL agent, by maximizing a reward function cantered on immediate access, tended to oversaturate high-crisis areas with resources, likely resulting in diminishing therapeutic returns and inefficient allocation. This suggests that future reward functions should incorporate stress decay, recovery state transitions, or diminishing marginal benefit curves to better align policy learning with human recovery dynamics [58]-[60]. The Mixed policy performed moderately across all metrics, offering a reasonable compromise between resource availability and effectiveness. Meanwhile, the Hotline-Focused strategy, while achieving broad coverage, failed to significantly reduce crisis or stress levels, reflecting the limited depth of low-intensity interventions in managing sustained psychological distress. Furthermore, the demographic analysis highlighted structural vulnerabilities among agents under 30, who experienced disproportionately higher crisis rates under all policies. The elevated stress under the RL strategy stems from its reward design, which maximizes coverage without penalizing oversaturation or diminishing returns. Agents may receive redundant interventions without sufficient recovery time, leading to chronic low-grade stress. Introducing a decay term or penalty for frequent but ineffective interventions may better align RL decisions with therapeutic recovery. Notably, the RL strategy minimized gender-based disparities more effectively than static approaches, suggesting that adaptive policies may be better suited to promoting equity, especially if trained with fairness-aware objectives. Collectively, these findings support a fundamental principle in mental health system design: effective crisis response requires not just scalable access, but context-aware, therapeutically rich interventions deployed with spatial and demographic sensitivity [61] [62]. Limitations of RL include difficulty in crafting reward functions that reflect real human recovery processes, sensitivity to feedback lags, and brittleness in unseen scenarios. Future work may benefit from hierarchical RL, human-in-the-loop reward shaping, or integrating fairness constraints to ensure broader applicability and ethical deployment. Simulation-based tools such as ABM combined with RL can offer a powerful, evidence-driven framework for optimizing policy strategies before real-world implementation [63]-[65].

5. Conclusion

This study demonstrates the power and flexibility of combining Agent-Based Modelling (ABM) with Reinforcement Learning (RL) to simulate, evaluate, and optimize mental health crisis response strategies at scale. By integrating a heterogeneous synthetic population, diverse intervention types, and both fixed and adaptive policy structures, the proposed framework enables granular experimentation with real-world relevance. Our results reveal that policy effectiveness cannot be solely measured by intervention volume. Strategies that prioritize targeted, therapeutically rich interventions—such as counselling-focused policies—consistently yield superior outcomes in reducing crisis rates, minimizing stress, and achieving higher efficiency scores [66] [67]. In contrast, policies optimized for maximum intervention coverage, such as those driven by reinforcement learning, may inadvertently lead to oversaturation, delivering frequent but potentially less impactful support, and sustaining elevated psychological tension in the population. The incorporation of demographic-specific outcome tracking and spatial heatmap analytics further demonstrates the framework's ability to uncover equity gaps and geographic inefficiencies in resource distribution. Such multidimensional insight is crucial for guiding evidence-based public health decisions, especially in resource-constrained or high-need settings [68] [69]. Importantly, the framework supports not only retrospective analysis but forward-looking policy simulation, enabling stakeholders to explore "what-if" scenarios, stress-test systems, and design fairer, more responsive mental health infrastructures. With further enhancement—such as reward shaping for RL agents, integration of real-world EHR or survey data, and modelling of budget constraints—this system has the potential to serve as a decision-support engine for mental health policy makers and urban planners [70] [71]. In conclusion, ABM+RL-based simulation offers a scalable, interpretable, and adaptable approach to the design and evaluation of mental health intervention strategies—one that bridges the gap between academic insight and operational policy impact.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Chua, B., Al-Ansi, A., Lee, M.J. and Han, H. (2020) Impact of Health Risk Perception on Avoidance of International Travel in the Wake of a Pandemic. *Current Issues in Tourism*, **24**, 985-1002. https://doi.org/10.1080/13683500.2020.1829570

[2] Rey, S.J. and Franklin, R.S. (2022) Introduction: Spatial Analysis and the Social Sciences in a Rapidly Changing Landscape. In: *Handbook of Spatial Analysis in the Social Sciences*, Edward Elgar Publishing, 11-20.

[3] LeNoble, C., Naranjo, A., Shoss, M. and Horan, K. (2023) Navigating a Context of Severe Uncertainty: The Effect of Industry Unsafety Signals on Employee Well-Being during the COVID-19 Crisis. *Occupational Health Science*, **7**, 707-743. https://doi.org/10.1007/s41542-023-00155-x

[4] Mollica, R., Cardozo, B.L., Osofsky, H., Raphael, B., Ager, A. and Salama, P. (2004) Mental Health in Complex Emergencies. *The Lancet*, **364**, 2058-2067. https://doi.org/10.1016/s0140-6736(04)17519-3

[5] North, C.S. and Pfefferbaum, B. (2013) Mental Health Response to Community Disasters: A Systematic Review. *JAMA*, **310**, 507-518.

[6] Li, G., Shi, W., Gao, X., Shi, X., Feng, X., Liang, D., *et al*. (2024) Mental Health and Psychosocial Interventions to Limit the Adverse Psychological Effects of Disasters and Emergencies in China: A Scoping Review. *The Lancet Regional Health—Western Pacific*, **45**, Article ID: 100580. https://doi.org/10.1016/j.lanwpc.2022.100580

[7] Bogdan, G.M., Seroka, A.M., Watson, J. and Johnson, M. (2007) Adapting Community Call Centers for Crisis Support: A Model for Home-Based Care and Monitoring. Agency for Healthcare Research and Quality.

[8] McEntire, D.A. (2021) Disaster Response and Recovery: Strategies and Tactics for Resilience. John Wiley & Sons.

[9] Afroogh, S., Mostafavi, A., Akbari, A., Pouresmaeil, Y., Goudarzi, S., Hajhosseini, F., *et al*. (2023) Embedded Ethics for Responsible Artificial Intelligence Systems (EE-RAIS) in Disaster Management: A Conceptual Model and Its Deployment. *AI and Ethics*, **4**, 1117-1141. https://doi.org/10.1007/s43681-023-00309-1

[10] Kirpalani, C. (2024) Technology-Driven Approaches to Enhance Disaster Response and Recovery. In: Kanga, S., *et al.*, Eds., Geospatial Technology for Natural Resource Management, Scrivener Publishing LLC, 25-81.

[11] Kirchner, T.R. and Shiffman, S. (2016) Spatio-Temporal Determinants of Mental Health and Well-Being: Advances in Geographically-Explicit Ecological Momentary Assessment (GEMA). *Social Psychiatry and Psychiatric Epidemiology*, **51**, 1211-1223. https://doi.org/10.1007/s00127-016-1277-5

[12] Wheaton, B. and Clarke, P. (2003) Space Meets Time: Integrating Temporal and Contextual Influences on Mental Health in Early Adulthood. *American Sociological Review*, **68**, 680-706. https://doi.org/10.1177/000312240306800502

[13] Nelson, B., McGorry, P.D., Wichers, M., Wigman, J.T.W. and Hartmann, J.A. (2017) Moving from Static to Dynamic Models of the Onset of Mental Disorder: A Review. *JAMA Psychiatry*, **74**, Article No. 528. https://doi.org/10.1001/jamapsychiatry.2017.0001

[14] Cummins, S., Curtis, S., Diez-Roux, A.V. and Macintyre, S. (2007) Understanding and Representing "Place" in Health Research: A Relational Approach. *Social Science & Medicine*, **65**, 1825-1838. https://doi.org/10.1016/j.socscimed.2007.05.036

[15] Patil, Y.M., Abraham, A.R., Chaubey, N.K., K., B. and Chidambaranathan, S. (2024) A Comparative Analysis of Machine Learning Techniques in Creating Virtual Replicas for Healthcare Simulations. In: Ponnusamy, S., *et al.*, Eds., *Harnessing AI and Digital Twin Technologies in Businesses*, IGI Global, 14-25. https://doi.org/10.4018/979-8-3693-3234-4.ch002

[16] Alhuwaydi, A. (2024) Exploring the Role of Artificial Intelligence in Mental Healthcare: Current Trends and Future Directions—A Narrative Review for a Comprehensive Insight. *Risk Management and Healthcare Policy*, **17**, 1339-1348. https://doi.org/10.2147/rmhp.s461562

[17] Koutsouleris, N., Hauser, T.U., Skvortsova, V. and De Choudhury, M. (2022) From Promise to Practice: Towards the Realisation of AI-Informed Mental Health Care. *The Lancet Digital Health*, **4**, e829-e840. https://doi.org/10.1016/s2589-7500(22)00153-4

[18] Karunakaran, M., Venkatachalam, C., Mahesh, T.R., Krishnan, B. and Nagaraj, S. (2024) Chapter 5. Machine Learning for Twinning the Human Body. In: Malviya, R., *et al.*, Eds., *Digital Transformation in Healthcare* 5.0, De Gruyter, 105-130. https://doi.org/10.1515/9783111398549-005

[19] Tracy, M., Cerdá, M. and Keyes, K.M. (2018) Agent-Based Modeling in Public Health: Current Applications and Future Directions. *Annual Review of Public Health*, **39**, 77-94. https://doi.org/10.1146/annurev-publhealth-040617-014317

[20] Silverman, B.G., Hanrahan, N., Bharathy, G., Gordon, K. and Johnson, D. (2015) A Systems Approach to Healthcare: Agent-Based Modeling, Community Mental Health, and Population Well-Being. *Artificial Intelligence in Medicine*, **63**, 61-71. https://doi.org/10.1016/j.artmed.2014.08.006

[21] 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.

[22] Yan, Y., Zhang, B. and Guo, J. (2016) An Adaptive Decision Making Approach Based on Reinforcement Learning for Self-Managed Cloud Applications. 2016 *IEEE International Conference on Web Services* (*ICWS*), San Francisco, 27 June-2 July 2016, 720-723. https://doi.org/10.1109/icws.2016.102

[23] Lewis, F.L., Vrabie, D. and Vamvoudakis, K.G. (2012) Reinforcement Learning and

Feedback Control: Using Natural Decision Methods to Design Optimal Adaptive Controllers. *IEEE Control Systems Magazine*, **32**, 76-105.

[24] Hussin, M., Asilah Wati Abdul Hamid, N. and Kasmiran, K.A. (2015) Improving Reliability in Resource Management through Adaptive Reinforcement Learning for Distributed Systems. *Journal of Parallel and Distributed Computing*, **75**, 93-100. https://doi.org/10.1016/j.jpdc.2014.10.001

[25] Marcus, N. and Stergiopoulos, V. (2022) Re-Examining Mental Health Crisis Intervention: A Rapid Review Comparing Outcomes across Police, Co-Responder and Non-Police Models. *Health & Social Care in the Community*, **30**, 1665-1679. https://doi.org/10.1111/hsc.13731

[26] Steadman, H.J., Deane, M.W., Borum, R. and Morrissey, J.P. (2000) Comparing Outcomes of Major Models of Police Responses to Mental Health Emergencies. *Psychiatric Services*, **51**, 645-649. https://doi.org/10.1176/appi.ps.51.5.645

[27] Paton, F., Wright, K., Ayre, N., Dare, C., Johnson, S., Lloyd-Evans, B., *et al.* (2016) Improving Outcomes for People in Mental Health Crisis: A Rapid Synthesis of the Evidence for Available Models of Care. *Health Technology Assessment*, **20**, 1-162. https://doi.org/10.3310/hta20030

[28] Garbarino, S. and Holland, J. (2009) Quantitative and Qualitative Methods in Impact Evaluation and Measuring Results.

[29] Boonekamp, P. (2006) Actual Interaction Effects between Policy Measures for Energy Efficiency—A Qualitative Matrix Method and Quantitative Simulation Results for Households. *Energy*, **31**, 2848-2873. https://doi.org/10.1016/j.energy.2006.01.004

[30] Allen, P., Pilar, M., Walsh-Bailey, C., Hooley, C., Mazzucca, S., Lewis, C.C., *et al.* (2020) Quantitative Measures of Health Policy Implementation Determinants and Outcomes: A Systematic Review. *Implementation Science*, **15**, Article No. 47. https://doi.org/10.1186/s13012-020-01007-w

[31] Pan, X.H., *et al.* (2022) Mate: Benchmarking Multi-Agent Reinforcement Learning in Distributed Target Coverage Control. *Advances in Neural Information Processing Systems* (*NeurIPS* 2021), 6-14 December 2021, 27862-27879.

[32] Aydemir, F. and Cetin, A. (2023) Multi-Agent Dynamic Area Coverage Based on Reinforcement Learning with Connected Agents. *Computer Systems Science and Engineering*, **45**, 215-230. https://doi.org/10.32604/csse.2023.031116

[33] Din, A., Ismail, M.Y., Shah, B., Babar, M., Ali, F. and Baig, S.U. (2022) A Deep Reinforcement Learning-Based Multi-Agent Area Coverage Control for Smart Agriculture. *Computers and Electrical Engineering*, **101**, Article ID: 108089. https://doi.org/10.1016/j.compeleceng.2022.108089

[34] Slade, M., Amering, M., Farkas, M., Hamilton, B., O'Hagan, M., Panther, G., *et al.* (2014) Uses and Abuses of Recovery: Implementing Recovery-Oriented Practices in Mental Health Systems. *World Psychiatry*, **13**, 12-20. https://doi.org/10.1002/wps.20084

[35] Duffy, P. and Baldwin, H. (2013) Recovery Post Treatment: Plans, Barriers and Motivators. *Substance Abuse Treatment*, *Prevention*, *and Policy*, **8**, 1-12. https://doi.org/10.1186/1747-597x-8-6

[36] Santangelo, P., Procter, N. and Fassett, D. (2017) Mental Health Nursing: Daring to Be Different, Special and Leading Recovery-Focused Care? *International Journal of Mental Health Nursing*, **27**, 258-266. https://doi.org/10.1111/inm.12316

[37] Glegg, S.M.N. and Levac, D.E. (2018) Barriers, Facilitators and Interventions to Support Virtual Reality Implementation in Rehabilitation: A Scoping Review. *PM&R*, **10**,

1237-1251. https://doi.org/10.1016/j.pmrj.2018.07.004

[38] Kunvardia, N. (2017) A Service Evaluation Study Exploring the Therapeutic Effectiveness of a Reiki Intervention in the Local Community of Cancer Patients. PhD Diss., Queen Margaret University.

[39] Swain, M. (1999) The New South Wales Drug Summit: Issues and Outcomes. NSW Parliamentary Library Research Service.

[40] Javeth, A., Salina, S. and Joshi, P. (2024) Dr. MT Bhatia Award Winner-Athar Javeth Aromatherapy for the Management of Cancer-Related. *Indian Journal of Palliative Care*, **30**, 121.

[41] Miller, N. (2015) Creating Opportunity after Crisis: Examining the Development of the Post-Earthquake Haitian Mental Health Care System. California Southern University.

[42] Berger, T. (2001) Agent-Based Spatial Models Applied to Agriculture: A Simulation Tool for Technology Diffusion, Resource Use Changes and Policy Analysis. *Agricultural Economics*, **25**, 245-260. https://doi.org/10.1016/s0169-5150(01)00082-2

[43] Bao, W., Gong, A., Zhang, T., Zhao, Y., Li, B. and Chen, S. (2023) Mapping Population Distribution with High Spatiotemporal Resolution in Beijing Using Baidu Heat Map Data. *Remote Sensing*, **15**, Article No. 458. https://doi.org/10.3390/rs15020458

[44] Ouda, E., Sleptchenko, A. and Simsekler, M.C.E. (2023) Comprehensive Review and Future Research Agenda on Discrete-Event Simulation and Agent-Based Simulation of Emergency Departments. *Simulation Modelling Practice and Theory*, **129**, Article ID: 102823. https://doi.org/10.1016/j.simpat.2023.102823

[45] Reininghaus, U., Reinhold, A.S., Priebe, S., Rauschenberg, C., Fleck, L., Schick, A., *et al.* (2024) Toward Equitable Interventions in Public Mental Health: A Review. *JAMA Psychiatry*, **81**, 1270-1275. https://doi.org/10.1001/jamapsychiatry.2024.3206

[46] Dodge, K.A., Prinstein, M.J., Evans, A.C., Ahuvia, I.L., Alvarez, K., Beidas, R.S., *et al.* (2024) Population Mental Health Science: Guiding Principles and Initial Agenda. *American Psychologist*, **79**, 805-823. https://doi.org/10.1037/amp0001334

[47] Fagan, A.A., Bumbarger, B.K., Barth, R.P., Bradshaw, C.P., Cooper, B.R., Supplee, L.H., *et al.* (2019) Scaling up Evidence-Based Interventions in US Public Systems to Prevent Behavioral Health Problems: Challenges and Opportunities. *Prevention Science*, **20**, 1147-1168. https://doi.org/10.1007/s11121-019-01048-8

[48] Thomson, K., Hillier-Brown, F., Todd, A., McNamara, C., Huijts, T. and Bambra, C. (2018) The Effects of Public Health Policies on Health Inequalities in High-Income Countries: An Umbrella Review. *BMC Public Health*, **18**, Article No. 869. https://doi.org/10.1186/s12889-018-5677-1

[49] Sangavi, C., Kollarmalil, R. and Abraham, S. (2025) Post-Mastectomy Wound Care—Need for an Empathetic Approach. *Psychology, Health & Medicine*. https://doi.org/10.1080/13548506.2025.2490229

[50] Okonkwo, O. (2019) Program Guide for Division 40 (Society Forclinical Neuropsychology) at the Annual Convention of the American Psychological Association, August 8-11, 2019; Chicago, IL. *The Clinical Neuropsychologist*, **33**, 1216-1348. https://doi.org/10.1080/13854046.2019.1628306

[51] Moriconi, C.B. (2003) A Systemic Treatment Program of Mindfulness Meditation for Fibromyalgia Patients and Their Partners. La Salle University.

[52] Rowe, J.B., Datta, D., Fiebach, C.J., Jaeggi, S.M., Liston, C., Luna, B., *et al.* (2024) Translating Prefrontal Cortex Insights to the Clinic and Society. In: Banich, M.T., *et al.*, Eds., *The Frontal Cortex*, The MIT Press, 319-360. https://doi.org/10.7551/mitpress/15679.003.0019

[53] Kurete, F. (2020) Enhancing Student Resilience through Access to Psychological Counselling Services in Selected Zimbabwean Polytechnics. PhD Diss., University of Pretoria (South Africa).

[54] Crowe, M., Eggleston, K., Douglas, K. and Porter, R.J. (2020) Effects of Psychotherapy on Comorbid Bipolar Disorder and Substance Use Disorder: A Systematic Review. *Bipolar Disorders*, **23**, 141-151. https://doi.org/10.1111/bdi.12971

[55] Kalusivalingam, A.K., Sharma, A., Patel, N. and Singh, V. (2020) Optimizing Industrial Systems through Deep Q-Networks and Proximal Policy Optimization in Reinforcement Learning. *International Journal of AI and ML*, **1**, 1-25.

[56] Gu, Y., Cheng, Y., Chen, C.L.P. and Wang, X. (2022) Proximal Policy Optimization with Policy Feedback. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, **52**, 4600-4610. https://doi.org/10.1109/tsmc.2021.3098451

[57] Schulman, J., Wolski, F., Dhariwal, P., Radford, A. and Klimov, O. (2017) Proximal Policy Optimization Algorithms.

[58] Memarzadeh, M. and Pozzi, M. (2019) Model-Free Reinforcement Learning with Model-Based Safe Exploration: Optimizing Adaptive Recovery Process of Infrastructure Systems. *Structural Safety*, **80**, 46-55. https://doi.org/10.1016/j.strusafe.2019.04.003

[59] Mark, M., Chehrazi, N., Liu, H. and Weber, T.A. (2022) Optimal Recovery of Unsecured Debt via Interpretable Reinforcement Learning. *Machine Learning with Applications*, **8**, Article ID: 100280. https://doi.org/10.1016/j.mlwa.2022.100280

[60] Nozhati, S., Sarkale, Y., Chong, E.K.P. and Ellingwood, B.R. (2020) Optimal Stochastic Dynamic Scheduling for Managing Community Recovery from Natural Hazards. *Reliability Engineering & System Safety*, **193**, Article ID: 106627. https://doi.org/10.1016/j.ress.2019.106627

[61] Zon, M., Ganesh, G., Deen, M.J. and Fang, Q. (2023) Context-Aware Medical Systems within Healthcare Environments: A Systematic Scoping Review to Identify Subdomains and Significant Medical Contexts. *International Journal of Environmental Research and Public Health*, **20**, Article No. 6399. https://doi.org/10.3390/ijerph20146399

[62] Thomas Craig, K.J., Morgan, L.C., Chen, C., Michie, S., Fusco, N., Snowdon, J.L., *et al*. (2020) Systematic Review of Context-Aware Digital Behavior Change Interventions to Improve Health. *Translational Behavioral Medicine*, **11**, 1037-1048. https://doi.org/10.1093/tbm/ibaa099

[63] Squazzoni, F. (2010) The Impact of Agent-Based Models in the Social Sciences after 15 Years of Incursion. *History of Economic Ideas*, **18**, 1000-1037.

[64] Lynch, C.J., Diallo, S.Y., Kavak, H. and Padilla, J.J. (2020) A Content Analysis-Based Approach to Explore Simulation Verification and Identify Its Current Challenges. *PLOS ONE*, **15**, e0232929. https://doi.org/10.1371/journal.pone.0232929

[65] Heath, B., Hill, R. and Ciarallo, F. (2009) A Survey of Agent-Based Modeling Practices (January 1998 to July 2008). *Journal of Artificial Societies and Social Simulation*, **12**, 9.

[66] Grant, R.W., Adams, A.S., Bayliss, E.A. and Heisler, M. (2013) Establishing Visit Priorities for Complex Patients: A Summary of the Literature and Conceptual Model to Guide Innovative Interventions. *Healthcare*, **1**, 117-122. https://doi.org/10.1016/j.hjdsi.2013.07.008

[67] Kazdin, A.E. (2017) Addressing the Treatment Gap: A Key Challenge for Extending Evidence-Based Psychosocial Interventions. *Behaviour Research and Therapy*, **88**, 7-18. https://doi.org/10.1016/j.brat.2016.06.004

[68] Ogundeko-Olugbami, O., Ogundeko, O., Lawan, M. and Foster, E. (2025) Harnessing Data for Impact: Transforming Public Health Interventions through Evidence-Based Decision-Making.

[69] Kashani, K.B., Awdishu, L., Bagshaw, S.M., Barreto, E.F., Claure-Del Granado, R., Evans, B.J., *et al*. (2023) Digital Health and Acute Kidney Injury: Consensus Report of the 27th Acute Disease Quality Initiative Workgroup. *Nature Reviews Nephrology*, **19**, 807-818. https://doi.org/10.1038/s41581-023-00744-7

[70] Luo, J.Y., Zhang, W.Z., Yuan, Y., *et al*. (2025) Large Language Model Agent: A Survey on Methodology, Applications and Challenges.

[71] Todd, J. and Stern, S. (2023) Towards More Nuanced Patient Management: Decomposing Readmission Risk with Survival Models. *The Operational Research Society's Annual Conference*, Bath, 12-14 September 2023, 157-158.