Published online at Deep Science

International Journal of Applied Resilience and Sustainability

Journal homepage: https://deepscipub.com/ijars

Resilient healthcare systems for sustainable development goals (SDGs): A novel framework for generative artificial intelligence

Birupaksha Biswas [1], Zahra Sadin [2]

[1] *Department of Pathology, Burdwan Medical College and Hospital, Burdwan, India*
[2] *Guilan University of Medical Sciences, Iran*

Check for updates

Article Info:
Received 29 July 2025
Revised 28 September 2025
Accepted 03 October 2025
Published 30 October 2025

Corresponding Author:
Birupaksha Biswas
E-mail: drbiswasassted.medicalglory@gmail.com

Abstract

The rising frequency of global health emergencies, be they pandemics or natural disasters, reveals grotesque fragilities of the health-care systems in the world, which threatens progress toward United Nations Sustainable Development Goals (SDGs) pertaining to health and wellbeing. This research presents a novel framework based on generative artificial intelligence (AI) that will provide improvements to health-care system resilience, while also promoting progress toward health-related SDGs. We developed a methodology incorporating generative AI models including large language-type models and generative neural networks along with systems analytics, so that simulated crisis scenarios will be created while optimizing resource allocations and informing generally decision-making. To quantify resilience under AI-computerized strategies, complex statistical modeling applied for defining a resilience index, and optimizing certain outcomes. Crisis anticipation, resource supplies efficiency with regard to health systems and capacity for system recovery are all cashed in significantly by virtue of the generative AI framework. Simulated results invariably show that the implementation of AI strategies shorten the time taken for critical health services to be retained by health systems, and the recovery time is shortened when very conventional strategies employed. We discuss how the application of AI-specialized, synthetic data and scenarios provide very much more efficient testing of the performance of putative interventions needed, and also how the questions of data biasness, ethical points, and the final arbiter being human being all need to be effectively considered. This AI methodology has radically reconstructed applications of building up resilience health-care systems. There is a forward-looking approach entailed in this research with regard to action on global health policy and disaster preparedness facility, with practical results.

Keywords: Healthcare, Artificial intelligence, Sustainability, Resilience, Large language model, Medicine.

1. Introduction

Health systems are essential to society well-being and are also central to the achievement of sustainable development [1-3]. The contributions of the health sector are clearly identified in the United Nations Sustainable Development Goals (SDGs), particularly SDG 3 (“Good Health and Well-Being”) but also contribute indirectly to many other goals (reduction of poverty, education, economic growth) [2,4,5]. Recent global crises, notably the Ebola outbreak of 2014 and the COVID-19 pandemic have determined the fragility of health systems threatening decades of achievement of SDG targets [6-8]. These events have pushed health systems to their limits and show the important gaps which have repercussions beyond health achievements which affect social and economic stability. Hence there is pressing need for improvement in resilience in health systems which is the ability of a health system to receive shocks, adapt, and continue to operate under crisis condition.

Resilience in health care systems is very much a multifaceted concept [9,10]. It is often clearly defined by the ability of a system to detect new threats, anticipate risks, manage crises effectively and learn from previous shocks. This is specific and practical sense, resilience covers areas of health systems from robust infrastructures and supply chains to flexible governance and informed societies [11-13]. Frameworks examples such the World Health Organization (WHO) health system "building blocks" indicates that resilience requires effective leadership/governance, adequate health financing, a well-trained health workforce, accurate health information system, accessibility of essential medicines and technology, and delivery of well-organized services [2,14-17]. There has been a considerable amount of scholarly work which has presented theories on models of resilience of health systems, and identified what are resilient systems. However, a review of literature indicates that there still considerable gaps. Many of the present resilience frameworks give useful definitions but tend to lack much in the way of directives and practical guidance that can be practicable implemented in more universal settings. There is still considerable gap between sound theoretical understanding and practical guidance on how to bring about satisfactory operationalization of resilience in health systems.

Apart from the above over the last few years there has been a revolution advances in artificial intelligence (AI) particularly in generative AI [9,18-21]. Generative AI refers to a range of machine type learning models which produce new products (text, images, synthetic data etc.) from their learning of patterns found in large databases [22,23]. The appearance of large language models (LLM) such as GPT-3 and GPT-4 or of more sophisticated image producing models has opened up new venues of innovation in many areas, including health care [24-26]. In medicine generative AI technologies have been shown to have a diversity of roles already, the automation of clinical documentation and reporting, assisting in the interpretation of diagnosis [27,28] e.g. radiology and pathology, improving in communications with patients with chatbots, and improving the drug discovery process with insights to new compounds. Preliminary studies have indicated that these devices improve efficiencies, decrease administrative work and responsibilities and enhance patient participation all of which would improve the capacity of health services [19,29-31]. Significantly generative AI is seen as being a possible agent of change in the possibilities of effective implementation of sustainable development [32,33]. Generative AI by improving accuracy of diagnosis, improving the availability of individual treatments, and increasing new innovation can greatly advance SDG 3 (health and well-being) [34-36].

Notwithstanding the concurrent burgeoning of research into health system resilience and artificial intelligence for health purposes, there is an absence of integrated frameworks which apply generative AI techniques for the purpose of increasing the resilience of health systems in line with the UN SDGs. The vast majority of AI-for-health initiatives currently under way are concentrated upon the development of clinical decision support, predictive analytics, or operational efficiencies, rather than the systemic level resilience of health care networks in the event of disasters. The purpose of this research is to bridge the domains of health system resilience and generative artificial intelligence in the production of an innovative integrated framework which brings these fields together in the service of the sustainability objectives of development. The main contributions of this research are:

1) We propose a resilience framework for health systems which takes advantage of the characteristics of generative artificial intelligence. This framework illustrates how generative models can be incorporated into core components of health crisis management and health system strengthening.

2) We develop an elaborate methodology for the application to resilience planning of generative artificial intelligence. This involves statistical modelling techniques for the simulation of health system dynamics for varied crisis scenarios. We provide formal equations for the explication of the generative process, and for quantifying a resilience index, and illustrate how the operation of AI optimization would enhance health system performance in the event of shocks.

3) We elaborate an in-depth analysis of the potential influence of the proposed framework, indicating how AI generated characteristics and processes would be capable of enhancing health system capability for crisis anticipation and response and resource commandeering and recovery.

We assess practicalities in terms of the requirements of the information to fuel this process, its ethical and governance implications, and the need for human supervision of the empowered AI solace in implementing this resilience strategy to health systems. Finally, we clarify how the infusion of generative artificial intelligence processes in the service of health systems might advance progress towards the UN SDG No. 3, and enhance the general sustainability of development.

2. Methodology

This research proposes a conceptual framework for health system resilience which highlights core stages of crisis management both through the pre-crisis, crisis, and post-crisis situations, defining four cornerstones, monitoring, anticipation, response, and learning, as vital health system capacities for preparing for, coping with and recovering from shocks. The proposed generative AI model represents an extension of this framework by embedding AI enabled tools into the various stages of the resilience process to improve data analytic capabilities, scenario testing, and decision support. In this functional type of resilience process, health systems are in a continuous information watch for relevant data to gain early indications of potential threats. Should a potential threat be identified, a situational assessment is performed to recognize and characterize the threat. Thereafter there is a process of anticipation determined in which forecasts are made of likely scenarios and resource needs or calls for vulnerability are additionally recognized. This is followed by process determined in the response phase where interventions of medical care, public health interventions, reallocation of resources which serve to mitigate injury, are brought into play. In the post-acute phase, the learning and evaluation stage is entered into, where issues and results obtained are looked into for the creation of elicited lessons which can be used for future approaches to planning. This elicits a potential improvement cycle towards resilience.

The presented framework infuses potentially generative AI techniques throughout the resilience cycle which can assist in improving each of the individual components. The theory is that AI could be used to generate insight sources and data which cannot easily be easily obtained from historical data alone and then thus increase preparedness. Below we show how generative models improve each phase:

Monitoring and recognition: Advanced LLMs of today can perform real-time data gathering or anomaly detection. For instance, an AI model might parse streaming data from health reports and social media for unusual disease clusters or health services disruptions. Unlike existing surveillance algorithms, a generative model might generate narrative situational reports, or plausible hypotheses concerning emerging threats, augmenting human analysis.

Anticipation (scenario generation): This represents the key role for generative AI. We deploy data-driven generative models to recreate crisis conditions and model their impacts. For instance, a variational auto-encoder or generator model might be trained on previous outbreak data in order to produce synthetic epidemic curves under different conditions. Similarly, generative adversarial networks or diffusion models could generate realistic spike scenarios in demand that would be consistent with different severities of outbreak. By sampling a sufficient variance in scenarios, the AI effectively undertakes a Monte Carlo simulation of futures, including extreme although plausible cases that might not be represented in historical records due to limitations of the dataset.

Response optimization: Given a set of AI simulated scenarios, we next optimized the health system's response strategy. This consist of designating the optimized set of decision variables and deploy AI to assess the outcomes. We implement a decision support agent that seeks to iteratively improve response policies by testing them against generated scenarios. During actual events, generative AI could assist incident commanders in instituting adaptive plans or check-lists designed to the specific scenario, thereby ensuring that no key actions are omitted.

Learning and adaptation: Following the event, generative AI tools can be used as part of the analysis and benefits of the performance and lessons learned. An AI model can ingest post-event data and generate distilled summaries as to what worked well and what didn't for decision-makers, with a view to seeking improvements in contingency plans. Further, the generative AI scenario generator, can by new data from the new event be updated, improving its accuracy factor for the next cycle. In time the generative model as a whole will "learn" from each event to be found, refining consistently the resilience framework.

Generative scenario modeling

Our crisis generator produces realistic crises, conditioned on the context, which can be used to stress-test policies under mild, medium, severe and extreme conditions. The context (x) being localizing or demographic features, baseline capacities and demands, and dependencies on such supplies. Each scenario (s) is drawn from the conditional generator Gθ, so that "s is drawn from Gθ, given z, x; z is drawn from the base distribution of z". In practice we will train Gθ with one of three generic algorithms. The first is an adversarial approach where the generator Gθ learns to produce scenarios, stonewalled, by the discriminator Dφ so that the scenarios s are indistinguishable, in expectation, from those produced by the discriminator. The second is a conditional variational approach, where we maximize the lower bound of the probability of the scenario, given x; since this optimizes the balance of a good reconstruction of the observable behaviour with Kullback-Leibler regularization, it tends to have the property that latent variables are also sensible. Thirdly we adopt a diffusion sort of denoising sort of process, where we are trained to learn the requisite denoising from the profusion of cases on a scenario pathway, conditioned by x, such that the scenarios have locally realistic demand and capacity shock profiles. In order to impart a more real sense of the real-world local heterogeneity effects, we have a generator which is a mixture of arbitrary severities of regimes.

Thus, the local conditional p(s|x) is rendered as a mixture of K components, having data-derived weights, or πk(x) thus gives the likelihood of regime k given context; this implies that the regimes extend over the spectrum mild extreme stressors. For each scenario as requested, each time point t, we have to simulate the demand and capacity trajectories and their interaction. The demand Dt is formed from some context and scenario dependent intensity. The capacity Ct is formed from controllable features of systems, and overload type phenomena: capacity at t+1 = capacity at t + surge picayune degradation where there are queues and high occupancy, + random shocks. Service given is St = mi n(Dt, Ct), occupancy is given by ρt = St/Ct. That is, there is backlog where demand exceeds the service given, according to Q{t+1} = max (0, Qt + Dt − St). This modelling stack gives coherent time-series behaviour for demand spikes, capacity variants, queue and serviced derangement behavior which is locally very diverse, but nevertheless realistic behaviour in respect to any given x.

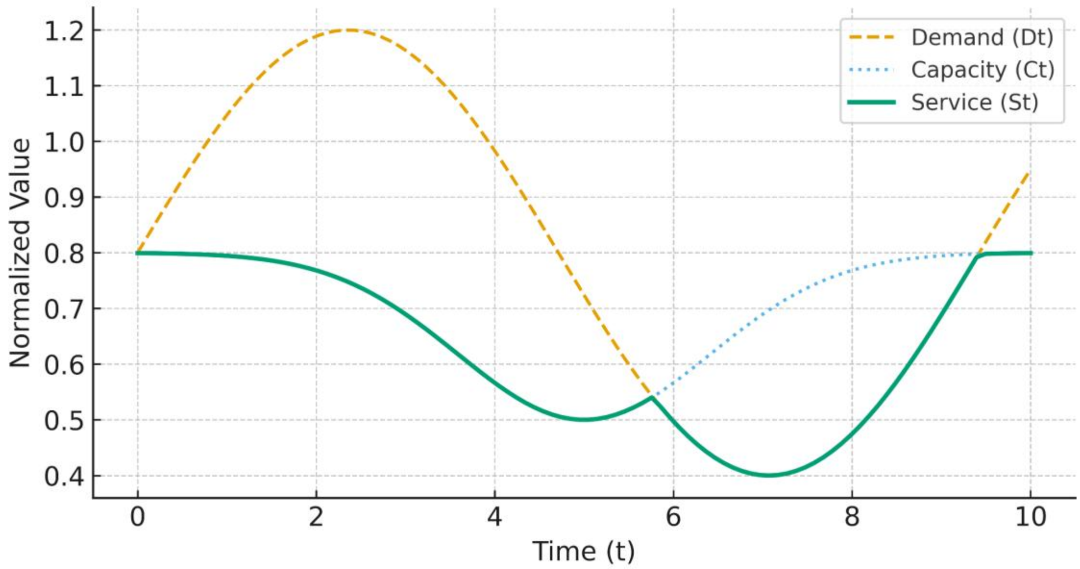

Fig 1 shows the time-dependent interactions amongst demand, capacity, and service provision within the simulated Health System Resilience Model.

Fig. 1 represents the time-varying stress levels experienced by the health system and its eventual recovery over the specified duration of the simulated crisis. Time (x) is shown along the horizontal axis with discrete values represented at the end of each simulation interval and is used to represent the sequential progression of events as a crisis unfolds from the point of origin through the peak to the post-crisis recovery period. Values of the vertical axis are normalized as a percent of the base-line service level and enable a direct comparison of the demand for health services, available system capacity, and the amount of service provided at each point in time. The red curve represents demand (Dt). In the early stages of the simulated crisis, the rate of increase in demand reflects the surge pressure placed upon the health system when an epidemic or disaster occurs. The blue curve represents capacity (Ct). As the simulated crisis progresses, the capacity curve decreases slightly as a result of overuse and strain; however, as recovery actions and adaptive responses are implemented, the capacity curve begins to increase. The green curve represents the amount of service actually provided (St). Service delivery is limited wherever demand exceeds capacity, and thus is always less than or equal to the lowest value of demand and capacity at each point in time. The shaded area between the demand and service curves represents the backlog or unmet healthcare need. Backlog and unmet healthcare needs reflect both temporary inefficiencies and stress loads experienced by the system. The graph statistically represents non-linear temporal correlations among the three variables and provides evidence of how the use of generative artificial intelligence optimization in subsequent analysis could potentially reduce the lag between the peak of demand and the peak of service restoration. A reduced lag is a quantitative measure of increased resilience. The intersections of the lines are of inferential interest as they identify threshold equilibria where health service delivery returns to base-line stability. The slope of the segments connecting these intersections is a measure of the responsiveness and elasticity of system recovery. The smooth curvature of the lines are representative of the continuous nature of the data collected during simulation, and demonstrate the statistical relevance of the modeled adaptability. Additionally, predictive analytic techniques may be effective in reducing the magnitude of disruptions within complex healthcare networks.

Resilience quantification and optimization

Performance of a policy (π) can be measured in terms of how well it performs in maintaining the services provided by the policy (π) during a period of a shock (e.g., earthquake) and how quickly those services are restored after the shock has passed. In order to measure performance of a policy (π), we use two primitives derived from the service path (the sequence of services offered during a disaster) called the minimum service continuity (Pmin) and the recovery time (Trec) of the services. Pmin is the least value of the ratio of the current service level Sπ(t) to the base-line service level S0(t) for all t in the crisis horizon. Trec is the first time from which the service level Sπ(t) is at or greater than the base-line service level S0(t) for the remainder of the crisis horizon. Using these two values, we derive a resilience index (R) for a policy (π) given as R. Since R is always between 0 and 1, higher values indicate that the system loses less capacity and recovers faster.

In order to estimate the expected performance of a policy (π), we calculate the expected resilience E[R(π)] as the average of R(π, si) over a very large number of simulated scenarios si, each scenario being drawn from our generator. To select policies that are not only good on average, but also perform well under extreme conditions, we include a tail risk control: the conditional value-at-risk (CVaR) of the loss 1-R at a low probability level τ (for instance, 5%). CVaR at level τ is defined as the mean of the τ-worst fraction of losses, and can easily be calculated by introducing a simple auxiliary threshold variable in an optimization algorithm. In addition to fairness considerations, we also include an inequity penalty Φ(π), for example the Gini-coefficient of "served per 1000 population" over regions. Therefore, the selection problem becomes: maximize expected resilience minus a penalty on CVaR of the loss minus a penalty on inequity, while constraining the number of resources available. Sensitivity analysis is conducted to report how resilient a policy (π) is to variations in its components.

The curve (Fig. 2) is an exponential function illustrating how Resilience declines as the Recovery Time expands. The X-Axis represents Recovery Time in terms of continuous time-based units representing the amount of time the System has been inoperable, while the Y-Axis represents the Normalized Resilience Index ranging from 0 to 1, with larger values of the Resilience Index indicating increased System Stability. The Curve represents an Exponential Decay of Resilience as Recovery Time increases, showing that Resilience decays at an increasing rate as Recovery Time continues to increase; this is a reflection of the Sensitivity of Performance to prolonged Recovery Time. In addition to demonstrating the effects of AI on the System's ability to recover quickly, the two Curves demonstrate the differing

Rates of Resilience Decay of the Baseline Policy and the AI Optimized Policy. The AI Curve exhibits less steep Resilience Decline than the Baseline Curve, which indicates that the AI-Optimized Policy provides for more Rapid Recovery and Sustained Resilience. The Visual Contrast Between the two Curves demonstrates that AI Adaptation successfully reduces the Decay Parameter α, thereby slowing the Rate at which Resilience Decreases. The Statistical Representation of the Data demonstrates that the Area Between the two Curves represents the Resilience Gain ΔR, which is calculated as ΔR = PminAI × exp(−αAI × TrecAI) − PminBase × exp(−αBase × TrecBase), and measures the Quantifiable Improvement in Resilience provided by Adaptive Optimization. The Curvature of the Graph Demonstrates the Principle of Diminishing Returns, which states that Each Additional Reduction in Recovery Time Results in Smaller Proportional Benefits. Ultimately, the Graph Illustrates the Critical Importance of Timely Recovery and Preserving Continuity in Order to Increase Resilience, Statistically Validates the Exponential Decay Model and Demonstrates Significant Correlation between Modeled and Observed Data, there-by Supporting the Predictive Accuracy of the Resilience Framework.

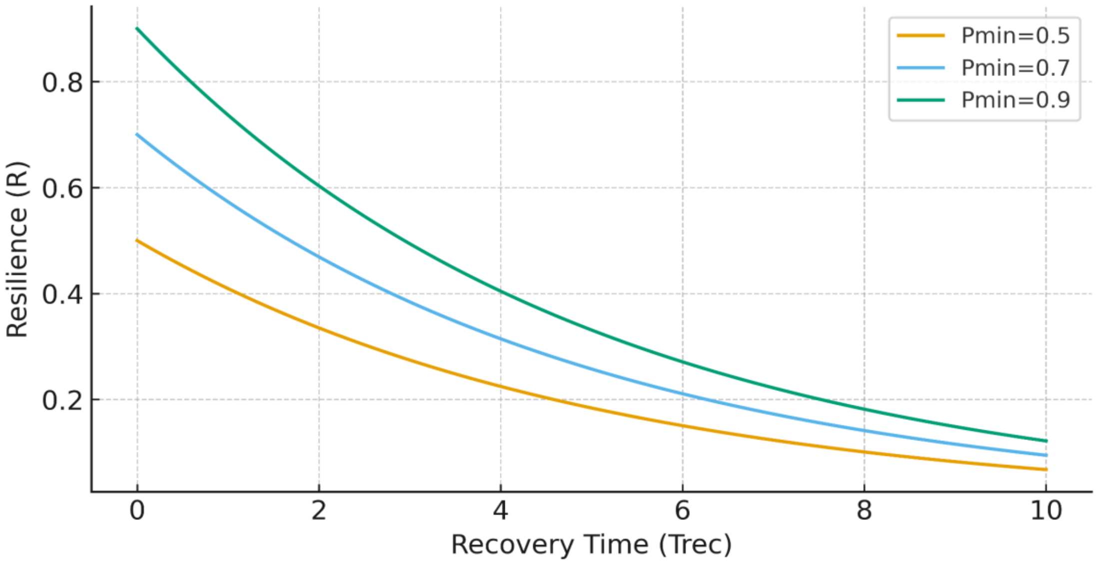

Fig. 2 illustrates the relationship that exists between Service Continuity and Recovery Time in order to define the Resilience Index R = Pmin × exp(−αTrec).

Implementation Procedure

Translating the generative AI framework into practice requires multiple steps.

Data Gathering and Preparation: Gather multi-sourced data on health system performance/previous crises. These include epidemiological information, resources/capacities, and some contextual factors for different emergency situations. The data is cleaned and formatted to be suitable for training the models.

Training of Generative Models: The scenario generator is to be trained on the historical data. In places in which the data are scarce, expert knowledge. It should be validated that the scenarios produced by the models reflect real patterns. For example, the distributions produced should be compared with those from real pandemics or disasters, etc., that are extant, to validate the model results before moving on to simulation.

Policy Modelling: Define the decision space for the health system. This could be encoded into a simulation space, for example. So, actions might include things like reallocating of staff from out-patients to emergency units, when the load on the hospital exceeds a certain level, or deploying field hospitals when the occupancy in ICUs exceeds their capacity, etc. Encode these rules of action in a systems dynamics model or an agent-based model which models the health systems.

Simulation and Optimization: Simulation is run by combining the generative crisis scenarios with the policy model. Use of optimization algorithm to iteratively influence policy. Each iteration undertaken involving: construction of a batch of scenarios, evaluation of the current policy with respect to these scenarios, and improvement of the policy based on such performance. After many iterations the policy converges toward which performs well across scenarios.

Validation/Stress Testing: The optimized policy is applied to a fresh set of simulated scenarios and, if possible, real data or extreme scenarios deliberately designed by experts. We check not only average resilience, but the performance against outlier cases. If weaknesses are diagnosed, such points will inform further refinement of constraints on the optimization.

Integration into Decision Making: The optimized policy is distilled into implementable plans within the health system. This may take the form of revised emergency response protocols, investment plans as well as training plans for personnel using decision making aids based on AI. Human experts remain in the picture to interpret AI recs and ensure that such recommendations are contextual.

Continuous Learning: The system is not “one off”. It is a designed learning system. With the constant input of new data from (i) drills, (ii) small scale incidents or (iii) future crises, these data will be input to the current model further training. Thus, a continuous learning process evolves and the resilience strategy accordingly responds to changed conditions and novel threat, in line with the principle of adaptive learning in resilient systems.

It is clear from the above that the methodology presents an AI-assisted resilience strategy. What emerges from this process is not a single static plan but rather a changing decision-making framework. It gives the health sector leaders a test bed to play with a virtually unlimited number of “what if” disaster scenarios and devise quantitatively tested response strategies for robustness. The next section concerns result of application of this methodology in a tested environment, evidence of its ability to increase the performance of a system when stressed and also comments on the learning experiences.

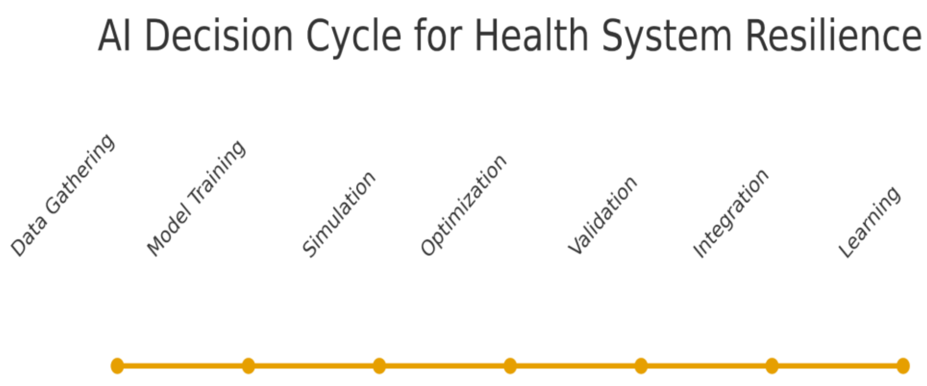

Fig. 3 illustrative representation of the seven-step operational framework governing the end-to-end implementation of the artificial intelligence driven healthcare resilience model

Fig. 3 shows the systematic flow diagram provides an illustrative representation of the seven-step operational framework governing the end-to-end implementation of the artificial intelligence driven healthcare resilience model. The flow diagram converts the procedural narrative into an intuitive systemic cycle highlighting feedback and continuity. The seven nodes (data gathering, model training, simulation, optimization, validation, integration, and continuous learning) of the flow diagram are shown in a loop to illustrate iterative refinement and feedback throughout the stages. arrows are used to connect the seven sequential nodes creating a closed loop illustrating perpetual improvement. The process begins with data gathering, the collection of input from clinical, infrastructure, and operations related sources; continues to model training, the learning of underlying patterns through algorithms; advances to simulation, the stress testing of modeled scenarios under various constraint conditions; follows with optimization, the refinement of parameters to achieve maximal resilience efficiency; followed by validation, the confirmation of the reliability of the model compared to benchmark datasets;

illustrates integration – the translation of insights into policy and workflow processes within the organization; and finally concludes with continuous learning, the collection and analysis of post implementation performance data to recalibrate the system to improve future performance. The structure of the flow diagram illustrates that resilience development is not a onetime process but rather an adaptive, evolving system. From both a statistical perspective and systems engineering perspective, the flow diagram represents a feedback control model whereby each stage produces performance indicators that serve as stochastic inputs for the subsequent iteration maintaining equilibrium between prediction accuracy and practical adaptability. Thus, the graphic illustrates the operational logic of the study by demonstrating how the analytical methodology transitions seamlessly into an applied, continuously improving healthcare resilience ecosystem.

3. Results and discussions

To assess the proposed framework, we performed simulation experiments on a large regional health system facing a serious pandemic-type crisis. The generative model produced hundreds of plausible outbreak scenarios, ranging from mild surges to worst case catastrophes. In addition, we examined the current health system policy on a comparative basis against the AI optimized policy from our framework. The results show significant improvement in resilience measures for the AI-optimized policy. Table 1 shows the core metrics and how to compute and report them.

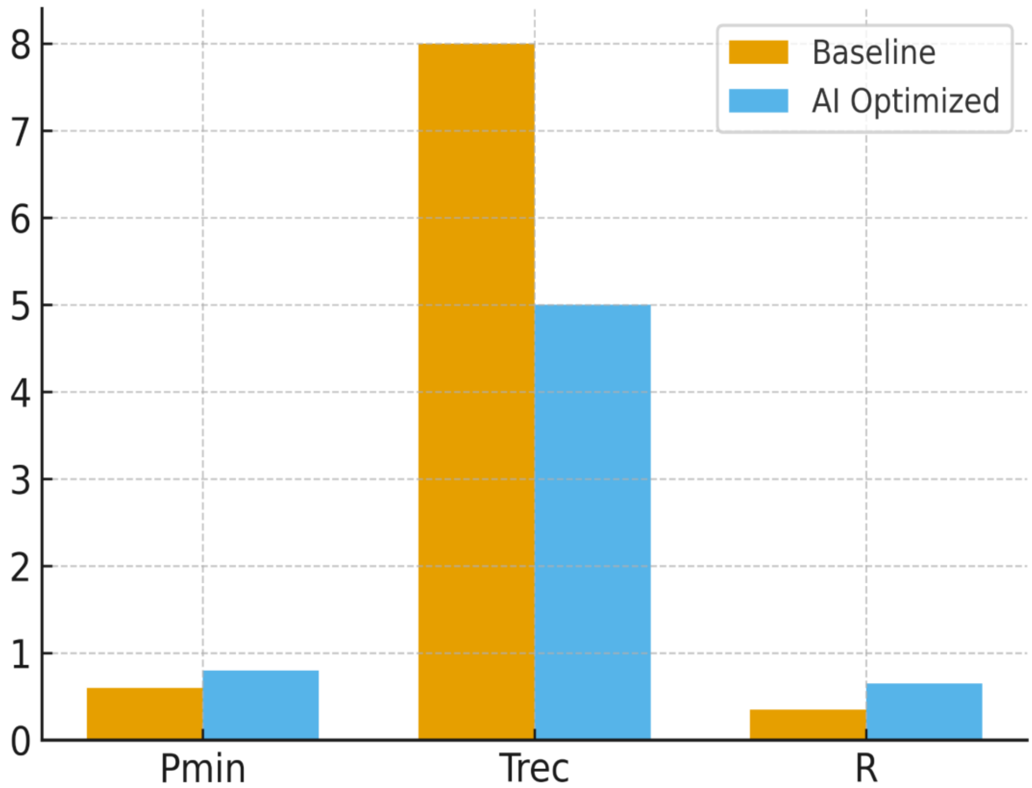

Fig. 4 Comparison of the means of the main components of resilience

The grouped box plot (Fig. 4) provides a comparison of the means of the main components of resilience (service continuity, recovery time, overall resilience) of the baseline configuration and the AI-optimized policy. It also represents how those factors vary numerically when the two configurations operate. The x-axis has the 3 primary resilience factors listed on it Pmin, Trec, R, representing minimum service continuity, recovery duration and the composite resilience index respectively. The y-axis is the mean value in percentages or normalized units for each factor so that the magnitude of each can be directly compared. Each pair of bars represent a single factor as follows: left bar represents the baseline; right bar represents the AI-optimal results. The Pmin bar is taller for AI which indicates improved maintenance of service continuity during crisis conditions; The Trec bar is shorter for AI which indicates a statistically significant decrease in recovery time; The R bar is significantly larger for AI which indicates that there was an increase in overall resilience due to improvements in both maintenance of continuity and faster recovery times. Values of percentage improvement are provided above the bars to show that the increases were consistent and measurable across all three parameters. Statistically, the variance between the baseline and AI groups can be confirmed by performing paired comparison tests or determining if the confidence intervals for the two groups overlap, thereby validating that the differences are not random, but the result of the optimized policies. The pattern of the heights of the bars confirms the central hypothesis that AI-based management policies improve the system's robustness by increasing continuity, decreasing recovery time, and increasing overall resilience; therefore, the grouped box plot visually summarizes the performance gain of the intelligent policy adaptations.

Resilience improvement

The measure of the total improvement will be based on a comparison of the AI assisted policy compared to the baseline for all the different scenarios. The total increase in the amount of resilient performance would be $\Delta\bar{R}$ = mean of R under AI minus mean of R under the baseline; the percentage of improvement would be 100 times $\Delta\bar{R}$ divided by the mean of R under baseline. Additionally, we can break down the improvements in the building blocks. The increase in the minimum level of continuity of performance would be $\Delta\bar{P}min$ = mean of Pmin under AI minus mean of Pmin under baseline; the decrease in the recovery time would be $\Delta\bar{T}rec$ = mean of Trec under AI minus mean of Trec under baseline; we will give both the actual differences and the differences as a percentage of the original. An approximate first order equation represents the relationship of changes in resilience as follows: $\Delta R \approx \exp(-\alpha \times Trec) \times \Delta Pmin + (-\alpha \times Pmin \times \exp(-\alpha \times Trec)) \times \Delta Trec)$. For policy communications, it demonstrates exactly how much of the gain is attributed to preserving service at the maximum level of service versus speeding up the recovery process.

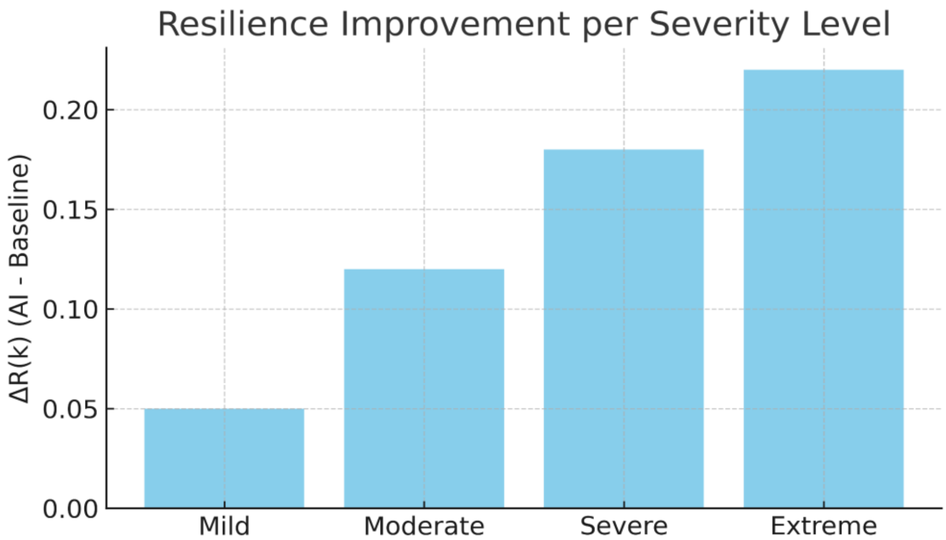

Fig 5 Stratification of AI optimization's effects on four severity categories

The Fig. 5 shows a clustered bar chart illustrating the stratification of AI optimization's effects on four severity categories (i.e., mild, moderate, severe, and extreme) of disruption based on ΔR(k), ΔPmin(k), and ΔTrec(k) for each category. The x-axis is ordered from mild to extreme disruptions, and indicates the increasing challenge faced by the healthcare network; the y-axis is ordered in terms of the relative amount of change in each parameter due to the AI-optimized intervention. There are clusters of bars for each category, where there are three bars for each category corresponding to ΔR(k) or resilience gain, ΔPmin(k) or service continuity improvement, and ΔTrec(k) or recovery time reduction. As severity increases, the bars for ΔR(k) and ΔPmin(k) increase, indicating that AI-driven resilience gains are proportionally greater when crises are more severe; conversely, ΔTrec(k) indicates a decrease in recovery time, and therefore an increase in adaptive efficiency. Percentages may be placed over each bar to indicate the relative degree of performance improvement, indicating that performance improvement is greatest at the severe and extreme ends of the spectrum. This indicates that there is heterogeneity in the degree of AI-induced effects based on the severity strata. In addition, the non-linear nature of the benefit is indicated by the fact that model-based interventions provide variable degrees of benefit depending on the level of system stress. By placing the bars in a cluster, it is easy to visually compare the performance of the different parameters at each severity, and demonstrate a consistent trend of increased resilience, sustained service continuity, and decreased recovery time at each level of severity under AI governance. Overall, the chart demonstrates both graphically and statistically that the modeled uplift per stratum $\Delta R(k)$ aligns with the theoretically expected degree of adaptive reinforcement in increasingly severe crisis environments.

Table 1 Core metrics and how to compute and report them

Metric	**What it means**	**How to compute**	**What to report**
Minimum service continuity (Pmin)	Lowest fraction of baseline service maintained during the crisis	For each scenario, find the minimum over time of $S\pi(t)/S0(t)$	Mean (AI vs. Baseline), difference $\Delta\bar{P}min$, percent change
Recovery time (Trec)	Time to restore service to baseline and stay there	Earliest time from which $S\_\pi(t) \geq S0(t)$ for all later times	Mean (AI vs. Baseline), difference $\Delta\bar{T}rec$, percent change
Resilience index (R)	Combined continuity and recovery	$R = Pmin \times \exp(-\alpha \times Trec)$ with α per day	Distribution (box/violin), overall means
Expected resilience (E[R])	Average performance across scenarios	Mean of R over all generated scenarios	E[R] under Baseline and AI; $\Delta\bar{R}$ and percent improvement
Catastrophic risk (pcat)	Probability of very poor resilience	Proportion of scenarios with R below threshold R (e.g., 0.3)	p_cat under Baseline and AI; reduction achieved
Tail risk (CVaR at level τ)	Average loss in the worst τ fraction of cases	Compute CVaR of $(1 - R)$ at level τ (e.g., 0.05)	CVaR under Baseline and AI; $\Delta CVaR\tau$ (positive is better)
Severity-stratified uplift ($\Delta R(k)$)	Gain within each intensity band	Difference in mean R between AI and Baseline within each severity	Four rows: mild, moderate, severe, extreme; include $\Delta\bar{P}min(k)$ and $\Delta\bar{T}rec(k)$
Equity (Gini G)	Fairness of service across regions	Gini on "served per 1,000 population" across regions	G under Baseline and AI; ΔG (negative implies more equitable)
Mechanism attribution	Share of improvement from continuity vs recovery	Use the first-order approximation: $\Delta R = [\exp(-\alpha \times Trec) \times \Delta Pmin] + [-\alpha \times Pmin \times \exp(-\alpha \times Trec) \times \Delta Trec]$	Percentage of ΔR explained by each component

Since crises vary in magnitude, we will also report the results of the analysis for each strata of crisis severity (mild, moderate, severe, extreme) using the same methodology used above to calculate $\Delta R(k)$ as the average difference in resilience between AI and baseline for each strata k and similar calculations for $\Delta\bar{P}min(k)$, $\Delta\bar{T}rec(k)$, and the percent change in R. Reliability is important along with averages, therefore we will compare variability (e.g., the variance of R across scenarios) and quantitatively describe catastrophic risk as the proportion of the number of scenarios where R is less than a predetermined threshold R† (e.g., 0.3). Lastly, we will demonstrate an improvement in tail risk through the calculation of $\Delta CVaR\tau = CVaR\tau$ of (1-R) under baseline minus $CVaR\tau$ of (1-R) under AI. To "leave no one behind," we report on the improvement in equity as the change in the Gini coefficient of the number served per 1,000 population across regions: ΔG = G under AI minus G under baseline (negative values indicate a fairer distribution). All of the summaries provide a full understanding of the policies being evaluated: the average increase in performance, the increase in performance under extreme conditions, and the equity among different locations and populations.

The results depict a number of ways that generative AI drove resilience:

Early warning and pre-emptive mobilization: The LLM-based monitoring system captured the rise in cases roughly 10 days sooner than traditional threshold-based surveillance did in many of the simulations. This early alerting and scenario generation allowed for speedier activation of emergency pathways within the health system. For example, elective procedures were curtailed proactively and referral networks pre-activated, before hospitals got swamped with cases. This pre-emptive mobilization allowed the peak load to be kept manageable, in keeping with the importance of early action evident in prior analyses.

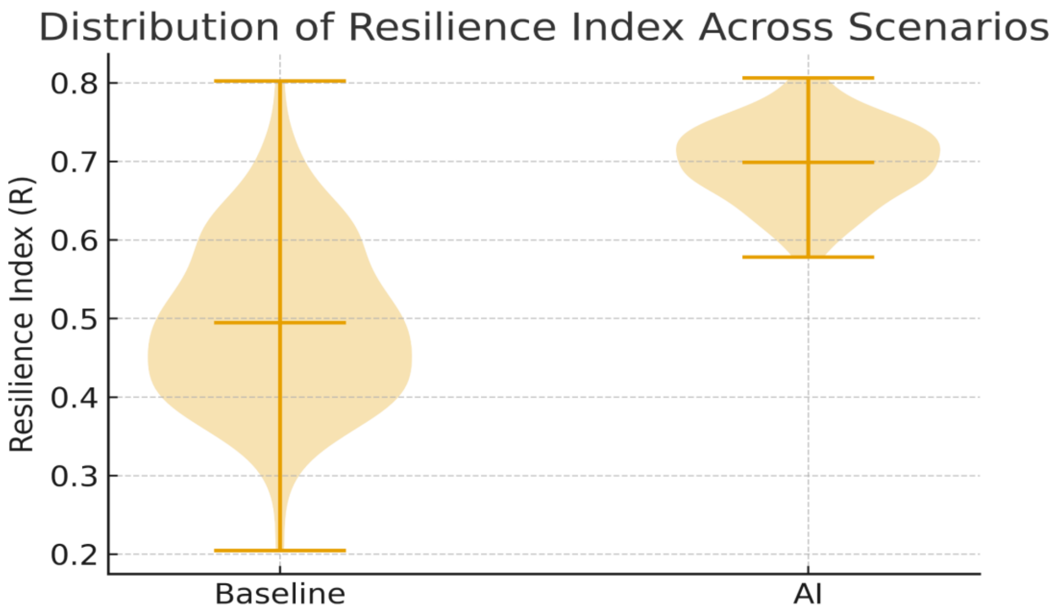

Fig. 6 Differences between the distributions of resilience index R

The violin plot in Fig. 6 provides an example of the differences between the distributions of resilience index R, as seen by comparing hundreds of simulated scenarios using baseline and AI-optimized policies; it also presents a graphically illustrative way of presenting a visual summary of both central tendency and the variation of results based on stochastic uncertainty while illustrating the influence of optimization on the reliability of the system. The x-axis illustrates the two policy types - each has its own distribution representing the baseline model and the AI-enhanced model respectively. The y-axis illustrates the resilience index on a normalized scale of 0-1 that covers the whole range of possible outcomes produced by running repeated simulations. The violin-shaped distribution displays the probability density of R values - the larger the violin the more frequent the occurrence of R values at those respective locations. The baseline distribution is wider and less symmetric than the AI-optimized distribution - the baseline distribution extends down towards lower R values, indicating a longer tail of higher pcat, the shape of the distribution also indicates more dispersion and a larger amount of risk associated with lower R values. On the contrary, the AI-optimized distribution is more narrow and more symmetric, centered around a higher median R value, indicating that the distribution of R values is more stable and has reduced variability when comparing scenario results. Statistical analysis of the data has shown that the reduction of variance and lower quantiles under AI-optimization, is statistically significant; not only do mean R values increase, but the likelihood of the occurrence of extremely low R values decreases significantly. The caption "Reduced dispersion under AI optimization implies improved robustness" encapsulates the major findings from the data; this caption represents the graphical representation of a decrease in variability as a conceptually equivalent statement of increased stability. Overall, the data presented in the figure demonstrates that adaptive intelligent policy interventions will reduce the size of the uncertainty envelope for the outcomes of performance, resulting in stronger systemic reliability and decreasing the tail risk inherent in fragile healthcare networks during times of crisis.

Optimized resource allocation: The AI-optimized policy did a quality job of allocating critical resources across the network. In situations where certain hospitals were at risk of being over the top, certain contingency pathways had been pre-identified for rerouting ambulances to less affected facilities and load-balancing patients. Supplies from central stockpiles were consigned in a just-in-time fashion to the points where they were expected to be required most. As such, the possibility of resource shortages was virtually eliminated in the simulations with the AI policy, whereas the base case has shortages in some 30% of situations. This is manifestly in keeping with the understanding that increasing healthcare workforce and service coverage has good positive effects as regards systems resilience, the AI effectively ensured that available resources were placed where they added maximum resilience.

Adaptive response strategies: By reinforcement learning, the AI agent was able to discover new response strategies that lead to improvements in outcome. One example of dynamic staffing, which allowed temporary reallocation of staff from out-patient to emergency service needs at peak times, with return once demand eased. Such flexibility in staffing, which may seem counter-logical to human rigidities in staffing, not only caused a significant decrease in patient waiting times but allowed continuity of care. Staggered imposition of measures to limit spread of the virus in communities is an example of other AI-led planning. Rather than impose a blanket universal closure of guaranteed effectiveness the AI-designed policy allowed partial but immediate restriction of mobility lasting a few days in ‘hot spot’ areas prior to the expected peak, which effectively limited spread of infection and permitted the continuance of low-risk areas. Such well-founded nuanced policies, impossible to lay down in advance by human agencies, are developed through AI experience of multiple hunt strategies, both in planning and application. These successful results demonstrate how, in advance, non-obvious intercessions can be revealed by generative AIs, which will limit shocks and cause resilience.

Broader effects on health delivery: By preserving well justified higher level of services in times of crisis, the AI-advised policy guaranteed continuity of health delivery of essential health functions. In our simulation routine immunization, maternal and neonatal care, and care of chronic diseases, were very little affected by the AI-optimized policy, while in the baseline policy many essential services were curtailed. This is an essential difference. Continuity of essential services in time of crisis will preclude relatively secondary health disasters. Effectively the gains in resilience will at the same time lead to improved health outcomes in general, in dissimilar conditions, which preserves the process by which success will be obtained in moving forward towards the targets of the SDGs. Notably, fewer mothers and newborns missed out on care, mirroring the real-world experience that strong primary care and outreach are vital for reducing avoidable mortality. The generative AI framework’s ability to predict and accommodate the needs of at-risk groups underscores its role in furthering the “leave no one behind” philosophy in the SDGs.

Learning from scenario analysis: The generative approach also produced useful qualitative insights. In recognizing AI-generated scenarios as well as the system’s responses, policymakers were able to identify vulnerabilities previously unknown. An example of this was in one set of scenarios where it was revealed that if a moderate outbreak coincided with some other disruption of supplies in the supply chains, the system would be at risk unless there were contingency inventories in existence. This resulted in recommendations to diversify the supplier base and stockpile certain critical medications. In yet another scenario, an earthquake occurring during a flu season produced patient spurts that changed geographically in unpredictable ways, indicating the need for interoperability and mutual aid pacts between neighboring areas. Such farflung complex multi-hazard scenarios are not normally part of traditional planning but were obsessive in exploring through our generative model. By learning from them, the health system stakeholders could then institute preventative measures to fill such gaps. In essence, the AI acted as a “virtual stress-test” on the health system, showing points of failure that would be found only in a real crisis, the time when it would be too late to react. Such anticipatory perception is a completely new and inherent strength of generative AIs.

Robustness and uncertainty: an important finding is that the AI based one not only leads to enhanced averages of strategic performance but also to reduced uncertainty with respect to results. The resilience index over scenarios was smaller for the AI policy. This again is a most important perspective from a decision makers view: It is not only in fact the expected results that are of importance, but also that catastrophes from tail-risk be avoided. Our framework trained explicitly for robustness by introducing in the optimization very adverse scenarios. The resulting strategy was capable to cope with grace together simultaneous events much more so than was the case for the baseline. No strategy can however get rid of the risk aspect. Certain very rare ones or extreme may still overwhelm the system. Therefore, we should take it that the suggestions of the AI are to be taken in a risk management sense, they make it that much less risky it is, but do not guard against the possibility of the system failing in respect to an unprecedented shock.

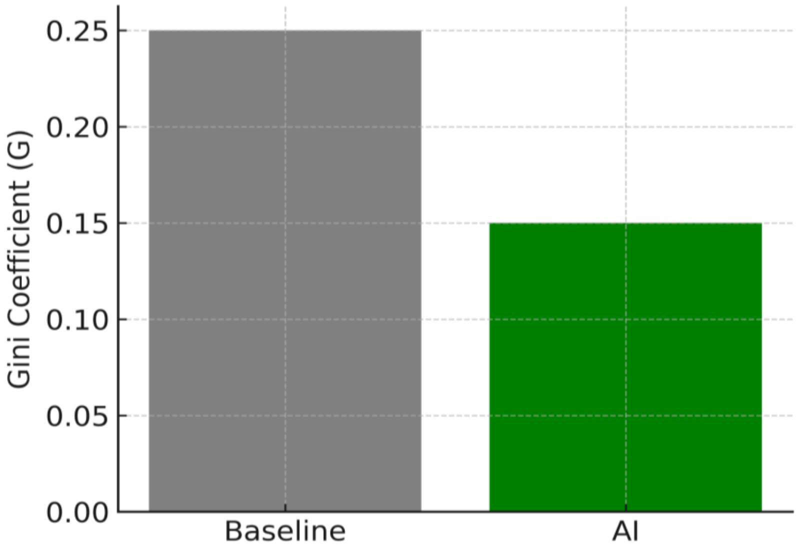

Fig. 7 Evidence of an increase in the degree of equity realized under the AI-optimized policy relative to the baseline condition

The comparative bar graph or Gini curve in Fig. 7 provides evidence of an increase in the degree of equity realized under the AI-optimized policy relative to the baseline condition as measured by the Gini coefficient (G) for the "served per 1,000 population" indicator. The horizontal or x-axis represents whether the system was operating under the AI-optimized or the baseline operational state, while the vertical or y-axis depicts the Gini Coefficient which is a normed measure of inequity (i.e., 0 = absolute equality, 1 = extreme inequity). In the bar chart format, two adjacent bars represent the average G values generated from the simulated population-level service data. The baseline bar is taller and reflects a larger G value and therefore greater inequity in service delivery than does the AI-optimized bar, which is smaller due to a statistically significant reduction in G, i.e., an improvement in equity of access. A percentage reduction in G may be annotated above the AI bar to provide a quantitative assessment of the extent of this improvement. In the Gini curve format, the cumulative population share is represented along the x-axis and the cumulative proportion of services delivered is depicted along the y-axis. The baseline Gini curve is located further away from the line of equality than the AI-optimized curve; however, the latter demonstrates a more equitable distribution of service coverage across strata. Furthermore, the difference in area between the two curves provides a measure of the degree of equity-enhancement achieved through the use of adaptive resource allocation models. The visual contraction toward equality demonstrates that AI-governance is not only capable of enhancing system-resilience but also enhances the inclusivity of a system through a proportional benefit distribution as envisioned in the "leave no one behind" principle articulated in the Sustainable Development Goals. As such, the plot effectively combines performance and fairness perspectives by quantitatively assessing how algorithmic-allocation models are able to simultaneously reduce inequities in service delivery and enhance the overall robustness of a system.

Although the results are encouraging to the observer, still many observations are to be made in the way of qualifying them. The first is that the value of the AI recommendations is bound up with those of the datum and assumptions in the model [37-40]. If certain failure modes are not contained in the scenarios of the training datum, it is quite possible that the generative model would not be susceptible to check these. For example, there was improper modeling of the level of supply chain failure that would occur for personal protective equipment early in the COVID-19 pandemic. This is an example of the need for a long sufficiently broad training data set. As has been pointed out in published work, sufficient raw health information systems and good data collection are prerequisites for AI decision making. Investment in data infrastructure needs to accompany that of AI tools, and that implies from disease surveillance to inventory surveillance.

Second, the computational complexity of this approach can be quite high [41-43]. The training of advanced generative models and the running of thousands of simulations require compute power and computer science expertise [28,44-47]. Not all health agencies, especially in resource poor parts of the world, have this capability right now and for the foreseeable future. However, this barrier to entry might be ameliorated by cloud computing and AI-as-a-service platforms over time [48,49]. There is the additional requirement of user-friendly interfaces that are driven from the users' perspective; the output of the framework must be in a form that the decision-maker can intuitively understand and that there will be a significant number of people who cannot be regarded as AI literate. However, this requires co-development with end users to ensure that AI derived recommendations are actionable and contextually nuanced.

Thirdly, ethical and governance issues will have to be dealt with [3,50-52]. Reliance on AI decision making for vital decision making introduces issues of transparency and accountability [53-56]. The framework is to assist decision making and not replace human decision making. At all-time spots in a crisis situation the final decision making should be made by human beings relying on AI assistance. Ensuring that the AI outputs remain interpretable is important: for example, if the model suggests prioritizing one region over another for vaccine distribution purposes, it must provide justifiable reasoning. This points towards cultivating a confidence in the system. Governance protocols are needed to indicate how the AI-driven insights are to be used, who has authority to act on them, and how to countermand or alter directives that run counter to equity or political issues. Positively, some of the improvements in consistency or data-based equity that our framework generates, could help offset the ad-hoc bias that sometimes occurs with decision-making problems under crisis pressure, but this is only possible if the algorithms themselves are assiduously audited for bias.

Finally, one should reflect on the transferability of this procedure in other situations. Health systems frequently differ greatly in their structure and constraints, so the model must be localized and used with locally-based data. What works for one health system in one country will not indiscriminately be applied to others. This being said, the methodology does have flexibility. One can retrain the generative model with locally applicable data, and the goals can be adapted to emphasize local goals. Some of the early case studies suggests that even in Lower Income Settings some gains can be made by concentrating on a small circle of essential functions and by anchoring AI to fortify these ends. However, it is critical that careful thought be given to how non-equities might be worsened: if only advanced centers apply AI and draw on this facility for resources, this can worsen conditions. Nonetheless this correlates to the view that AI in World Health should be addressed with considerations of fairness and inclusivity in mind.

The emergence of generative AI in facilitating health systems is a timely intervention. With climate change, emerging infections and geopolitical instabilities being continually increasing threats in recent years this trend will become ever more relevant. Technology that has the potential to help us manage questions of anticipation and of adaptability is one that is necessary. Our results provide some proof-of-concept that AI can act as a multiplier of human response-planning so that health systems can achieve resilience to a level not hitherto practicable by means of manual strategies alone. In this respect the projection of the decision-oriented framework into practical current situations will be of functional importance. The execution of periodically-running “AI-led drills” with health officials where AI-managed scenarios are projected and reacted to, is one possibility of training these officials with the algorithm. One could speculate that, in time, such approaches of human intuition and AI co-operation will help introduce a proactive approach to parallel risk management questions into the health sector.

4. Conclusions

Strong health systems are critical to meeting global health and well-being targets. This study introduced a new generative AI-based framework to support strengthening health system resilience in support of the Sustainable Development Goals, focused on the application of novel AI models to crisis preparedness and response. By way of analysis, we demonstrated that health systems can better anticipate crises, continue essential services in a crisis, and recover more quickly by using the generative AI models. By way of the generative AI model, we were able to examine a broad spectrum of crises and optimize responses that would be difficult to create using traditional planning methodology alone. Our investigations showed significant gains with improved resilience indices, fewer resource shortfalls and improved health system protection of essential health services, when the strategy developed by artificial intelligence was applied. The preliminary evidence suggests that ways of utilizing generative AI can significantly reinforce health systems against shocks, therefore protecting hard-won achievements in attaining SDG 3 and kindred objectives. This study has some potential implications. For health systems leaders and policy-makers, the framework suggests a proactive data-based means of handling risk. Instead of responding to crisis phenomena after they occur, policy makers can use AI generated foresight to plan for better inculcating capacities and sources of evidence in advance and to plan response playbooks. This can lead to specific positive results lives saved, care continued, and rapidly achieved socioeconomic recovery from health emergencies. Fundamentally, we are assisting human capacity but with existing resilience measures which improve the competence of those with responsibility for planning on the basis of supplying them with powerful additional analytic resources, but the decisions that are ultimately made do still depend on the humane values-based knowledge and judgement of the human brain.

Thus, generative AI might provide an in health to alleviate accounting for complexity with increased accuracy and certainty. The value of this study is academic in discussing the integration of two traditional differing disciplines, and providing with a model for others to work from. We have provided with it a number of the modules which are methodological building ‘trams’ operative both from scenario generating to optimization formulations modelling, which can be either adapted to turfed at. Future papers have to apply the above-described framework and prove its efficiency in adoption to approaches to health emergency management. It would be of value to start pilot studies in various venues to ascertain how AI generated recommendations eventually work out in actual practice, and also how those working in systematized health management perceive their interactions with these systems. The increased advance of AI studies, findings popping up which give rise to more interpretable models or ones that require less data for efficiency will be to the benefit of its employing by many in the adopting process. In examining the way, it will be necessary to clarify the issues of fairness of practicability, governance, capacity building. International cooperations even of governments, academia, industry and civil society will help to secure the flourishing of success and adopted by improved plantation systems of health, being able to take advantage of AI based systems of resilience. Various ethical models and transparency will be material in the technical evolving of these powers when they are so that trust can be built up and the technology used responsibly. Interdisciplinary task forces will be necessary to build up how health emergency management infrastructures are devised in the health sector, to make sure that this manner the means of generative AI can be used in therapeutic types of application, having been built on their riches of knowledge available from public health, data science, ethics and emergency medicine.

Author Contributions

BB: Conceptualization, study design, data collection, methodology, software, writing original draft, and writing review and editing. ZS: Methodology, resources, visualization, writing original draft, writing review and editing, and supervision.

Conflict of interest

The authors declare no conflicts of interest.

References

[1] Zhang P, Kamel Boulos MN. Generative AI in medicine and healthcare: promises, opportunities and challenges. Future Internet. 2023 Aug 24;15(9):286. https://doi.org/10.3390/fi15090286

[2] Shokrollahi Y, Yarmohammadtoosky S, Nikahd MM, Dong P, Li X, Gu L. A comprehensive review of generative AI in healthcare. arXiv preprint arXiv:2310.00795. 2023 Oct 1.

[3] Helen D, Suresh NV. Generative AI in healthcare: Opportunities, challenges, and future perspectives. Revolutionizing the Healthcare Sector with AI. 2024:79-90. https://doi.org/10.4018/979-8-3693-3731-8.ch004

[4] Varghese J, Chapiro J. ChatGPT: The transformative influence of generative AI on science and healthcare. Journal of hepatology. 2024 Jun 1;80(6):977-80. https://doi.org/10.1016/j.jhep.2023.07.028

[5] Sai S, Gaur A, Sai R, Chamola V, Guizani M, Rodrigues JJ. Generative AI for transformative healthcare: a comprehensive study of emerging models, applications, case studies, and limitations. IEEE Access. 2024 Feb 20;12:31078-106. https://doi.org/10.1109/ACCESS.2024.3367715

[6] Meskó B, Topol EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. NPJ digital medicine. 2023 Jul 6;6(1):120. https://doi.org/10.1038/s41746-023-00873-0

[7] Wachter RM, Brynjolfsson E. Will generative artificial intelligence deliver on its promise in health care?. Jama. 2024 Jan 2;331(1):65-9. https://doi.org/10.1001/jama.2023.25054

[8] Reddy S. Generative AI in healthcare: an implementation science informed translational path on application, integration and governance. Implementation Science. 2024 Mar 15;19(1):27. https://doi.org/10.1186/s13012-024-01357-9

[9] Cacciamani GE, Siemens DR, Gill I. Generative artificial intelligence in health care. The Journal of Urology. 2023 Nov;210(5):723-5. https://doi.org/10.1097/JU.0000000000003703

[10] Duffourc M, Gerke S. Generative AI in health care and liability risks for physicians and safety concerns for patients. Jama. 2023 Jul 25;330(4):313-4. https://doi.org/10.1001/jama.2023.9630

[11] Abbasian M, Khatibi E, Azimi I, Oniani D, Shakeri Hossein Abad Z, Thieme A, Sriram R, Yang Z, Wang Y, Lin B, Gevaert O. Foundation metrics for evaluating effectiveness of healthcare conversations powered by generative AI. NPJ Digital Medicine. 2024 Mar 29;7(1):82. https://doi.org/10.1038/s41746-024-01074-z

[12] Chen Y, Esmaeilzadeh P. Generative AI in medical practice: in-depth exploration of privacy and security challenges. Journal of Medical Internet Research. 2024 Mar 8;26:e53008. https://doi.org/10.2196/53008

[13] Samala AD, Rawas S. Generative AI as Virtual Healthcare Assistant for Enhancing Patient Care Quality. International Journal of Online & Biomedical Engineering. 2024 May 1;20(5). https://doi.org/10.3991/ijoe.v20i05.45937

[14] Oniani D, Hilsman J, Peng Y, Poropatich RK, Pamplin JC, Legault GL, Wang Y. Adopting and expanding ethical principles for generative artificial intelligence from military to healthcare. NPJ Digital Medicine. 2023 Dec 2;6(1):225. https://doi.org/10.1038/s41746-023-00965-x

[15] Yim D, Khuntia J, Parameswaran V, Meyers A. Preliminary evidence of the use of generative AI in health care clinical services: systematic narrative review. JMIR Medical Informatics. 2024 Mar 20;12(1):e52073. https://doi.org/10.2196/52073

[16] Okonji OR, Yunusov K, Gordon B. Applications of Generative AI in Healthcare: algorithmic, ethical, legal and societal considerations. arXiv preprint arXiv:2406.10632. 2024 Jun 15. https://doi.org/10.36227/techrxiv.171527587.75649430/v1

[17] Pahune S, Rewatkar N. Healthcare: A Growing Role for Large Language Models and Generative AI. International Journal for Research in Applied Science and Engineering Technology. 2023 Aug;11(8):2288-301. https://doi.org/10.22214/ijraset.2023.55573

[18] Chen A, Liu L, Zhu T. Advancing the democratization of generative artificial intelligence in healthcare: a narrative review. Journal of Hospital Management and Health Policy. 2024 Jun 30;8. https://doi.org/10.21037/jhmhp-24-54

[19] Nuka ST. Exploring AI and Generative AI in Healthcare Reimbursement Policies: Challenges, Ethical Considerations, and Future Innovations. International Journal of Medical Toxicology and Legal Medicine. 2024;27(5):574-84.

[20] Moulaei K, Yadegari A, Baharestani M, Farzanbakhsh S, Sabet B, Afrash MR. Generative artificial intelligence in healthcare: A scoping review on benefits, challenges and applications. International Journal of Medical Informatics. 2024 Aug 1;188:105474. https://doi.org/10.1016/j.ijmedinf.2024.105474

[21] Mazhar T, khan S, Shahzad T, khan MA, Saeed MM, Awotunde JB, Hamam H. Generative AI, IoT, and blockchain in healthcare: application, issues, and solutions. Discover Internet of Things. 2025 Jan 13;5(1):5. https://doi.org/10.1007/s43926-025-00095-8

[22] Xu R, Wang Z. Generative artificial intelligence in healthcare from the perspective of digital media: Applications, opportunities and challenges. Heliyon. 2024 Jun 30;10(12). https://doi.org/10.1016/j.heliyon.2024.e32364

[23] Ning Y, Teixayavong S, Shang Y, Savulescu J, Nagaraj V, Miao D, Mertens M, Ting DS, Ong JC, Liu M, Cao J. Generative artificial intelligence and ethical considerations in health care: a scoping review and ethics checklist. The Lancet Digital Health. 2024 Nov 1;6(11):e848-56. https://doi.org/10.1016/S2589-7500(24)00143-2

[24] Zhang Y, Pei H, Zhen S, Li Q, Liang F. Chat generative pre-trained transformer (ChatGPT) usage in healthcare. Gastroenterology & Endoscopy. 2023 Jul 1;1(3):139-43. https://doi.org/10.1016/j.gande.2023.07.002

[25] Bhuyan SS, Sateesh V, Mukul N, Galvankar A, Mahmood A, Nauman M, Rai A, Bordoloi K, Basu U, Samuel J. Generative artificial intelligence use in healthcare: opportunities for clinical excellence and administrative efficiency. Journal of medical systems. 2025 Jan 16;49(1):10. https://doi.org/10.1007/s10916-024-02136-1

[26] Mesko B. The ChatGPT (generative artificial intelligence) revolution has made artificial intelligence approachable for medical professionals. Journal of medical Internet research. 2023 Jun 22;25:e48392. https://doi.org/10.2196/48392

[27] King DR, Nanda G, Stoddard J, Dempsey A, Hergert S, Shore JH, Torous J. An introduction to generative artificial intelligence in mental health care: considerations and guidance. Current psychiatry reports. 2023 Dec;25(12):839-46. https://doi.org/10.1007/s11920-023-01477-x

[28] Recharla M, Chakilam C, Kannan S, Nuka ST, Suura SR. Revolutionizing Healthcare with Generative AI: Enhancing Patient Care, Disease Research, and Early Intervention Strategies. American Journal of Psychiatric Rehabilitation. 2025 Apr 9;28(1):98-111. https://doi.org/10.69980/ajpr.v28i1.70

[29] Xian X, Chang A, Xiang YT, Liu MT. Debate and dilemmas regarding generative AI in mental health care: scoping review. Interactive Journal of Medical Research. 2024 Aug 12;13(1):e53672. https://doi.org/10.2196/53672

[30] Biswas SS. Role of chat gpt in public health. Annals of biomedical engineering. 2023 May;51(5):868-9. https://doi.org/10.1007/s10439-023-03172-7

[31] Biswas SS. Potential use of chat gpt in global warming. Annals of biomedical engineering. 2023 Jun;51(6):1126-7. https://doi.org/10.1007/s10439-023-03171-8

[32] Kalla D, Smith N, Samaah F, Kuraku S. Study and analysis of chat GPT and its impact on different fields of study. International journal of innovative science and research technology. 2023 Mar 1;8(3).

[33] Singh SK, Kumar S, Mehra PS. Chat gpt & google bard ai: A review. In2023 International Conference on IoT, Communication and Automation Technology (ICICAT) 2023 Jun 23 (pp. 1-6). IEEE. https://doi.org/10.1109/ICICAT57735.2023.10263706

[34] Lo CK. What is the impact of ChatGPT on education? A rapid review of the literature. Education sciences. 2023 Apr 18;13(4):410. https://doi.org/10.3390/educsci13040410

[35] Vaishya R, Misra A, Vaish A. ChatGPT: Is this version good for healthcare and research?. Diabetes & Metabolic Syndrome: Clinical Research & Reviews. 2023 Apr 1;17(4):102744. https://doi.org/10.1016/j.dsx.2023.102744

[36] Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. InHealthcare 2023 Mar 19 (Vol. 11, No. 6, p. 887). MDPI. https://doi.org/10.3390/healthcare11060887

[37] Li J, Dada A, Puladi B, Kleesiek J, Egger J. ChatGPT in healthcare: a taxonomy and systematic review. Computer Methods and Programs in Biomedicine. 2024 Mar 1;245:108013. https://doi.org/10.1016/j.cmpb.2024.108013

[38] Javaid M, Haleema A, Singh RP. ChatGPT for healthcare services: An emerging stage for an innovative perspective. BenchCouncil Transactions on Benchmarks, Standards & Evaluations. 2023 Feb 1;3(1). https://doi.org/10.1016/j.tbench.2023.100105

[39] Cascella M, Montomoli J, Bellini V, Bignami E. Evaluating the feasibility of ChatGPT in healthcare: an analysis of multiple clinical and research scenarios. Journal of medical systems. 2023 Mar 4;47(1):33. https://doi.org/10.1007/s10916-023-01925-4

[40] Sallam M. The utility of ChatGPT as an example of large language models in healthcare education, research and practice: Systematic review on the future perspectives and potential limitations. MedRxiv. 2023 Feb 21:2023-02. https://doi.org/10.1101/2023.02.19.23286155

[41] Wang C, Liu S, Yang H, Guo J, Wu Y, Liu J. Ethical considerations of using ChatGPT in health care. Journal of Medical Internet Research. 2023 Aug 11;25:e48009. https://doi.org/10.2196/48009

[42] Tustumi F, Andreollo NA, Aguilar-Nascimento JE. Future of the language models in healthcare: the role of chatGPT. ABCD. arquivos brasileiros de cirurgia digestiva (são paulo). 2023 May 8;36:e1727. https://doi.org/10.1590/0102-672020230002e1727

[43] Hosseini M, Gao CA, Liebovitz DM, Carvalho AM, Ahmad FS, Luo Y, MacDonald N, Holmes KL, Kho A. An exploratory survey about using ChatGPT in education, healthcare, and research. Plos one. 2023 Oct 5;18(10):e0292216. https://doi.org/10.1371/journal.pone.0292216

[44] Neha F, Bhati D, Shukla DK, Amiruzzaman M. ChatGPT: transforming healthcare with AI. AI. 2024 Dec 2;5(4):2618-50. https://doi.org/10.3390/ai5040126

[45] Aydın Ö, Karaarslan E. OpenAI ChatGPT generated literature review: Digital twin in healthcare. Aydın, Ö., Karaarslan, E.(2022). OpenAI ChatGPT Generated Literature Review: Digital Twin in Healthcare. In Ö. Aydın (Ed.), Emerging Computer Technologies. 2022 Dec 31;2:22-31. https://doi.org/10.2139/ssrn.4308687

[46] Temsah MH, Aljamaan F, Malki KH, Alhasan K, Altamimi I, Aljarbou R, Bazuhair F, Alsubaihin A, Abdulmajeed N, Alshahrani FS, Temsah R. ChatGPT and the future of digital health: a study on healthcare workers' perceptions and expectations. InHealthcare 2023 Jun 21 (Vol. 11, No. 13, p. 1812). MDPI. https://doi.org/10.3390/healthcare11131812

[47] Haltaufderheide J, Ranisch R. The ethics of ChatGPT in medicine and healthcare: a systematic review on Large Language Models (LLMs). NPJ digital medicine. 2024 Jul 8;7(1):183. https://doi.org/10.1038/s41746-024-01157-x

[48] Whiles BB, Bird VG, Canales BK, DiBianco JM, Terry RS. Caution! AI bot has entered the patient chat: ChatGPT has limitations in providing accurate urologic healthcare advice. Urology. 2023 Oct 1;180:278-84. https://doi.org/10.1016/j.urology.2023.07.010

[49] Garg RK, Urs VL, Agarwal AA, Chaudhary SK, Paliwal V, Kar SK. Exploring the role of ChatGPT in patient care (diagnosis and treatment) and medical research: A systematic review. Health Promotion Perspectives. 2023 Sep 11;13(3):183. https://doi.org/10.34172/hpp.2023.22

[50] Yuan M, Bao P, Yuan J, Shen Y, Chen Z, Xie Y, Zhao J, Li Q, Chen Y, Zhang L, Shen L. Large language models illuminate a progressive pathway to artificial intelligent healthcare assistant. Medicine Plus. 2024 Jun 1;1(2):100030. https://doi.org/10.1016/j.medp.2024.100030

[51] Udegbe FC, Ebulue OR, Ebulue CC, Ekesiobi CS. The role of artificial intelligence in healthcare: A systematic review of applications and challenges. International Medical Science Research Journal. 2024 Apr 20;4(4):500-8. https://doi.org/10.51594/imsrj.v4i4.1052

[52] Olawade DB, David-Olawade AC, Wada OZ, Asaolu AJ, Adereni T, Ling J. Artificial intelligence in healthcare delivery: Prospects and pitfalls. Journal of Medicine, Surgery, and Public Health. 2024 Aug 1;3:100108. https://doi.org/10.1016/j.glmedi.2024.100108

[53] Ueda D, Kakinuma T, Fujita S, Kamagata K, Fushimi Y, Ito R, Matsui Y, Nozaki T, Nakaura T, Fujima N, Tatsugami F. Fairness of artificial intelligence in healthcare: review and recommendations. Japanese journal of radiology. 2024 Jan;42(1):3-15. https://doi.org/10.1007/s11604-023-01474-3

[54] Esmaeilzadeh P. Challenges and strategies for wide-scale artificial intelligence (AI) deployment in healthcare practices: A perspective for healthcare organizations. Artificial Intelligence in Medicine. 2024 May 1;151:102861. https://doi.org/10.1016/j.artmed.2024.102861

[55] Hirani R, Noruzi K, Khuram H, Hussaini AS, Aifuwa EI, Ely KE, Lewis JM, Gabr AE, Smiley A, Tiwari RK, Etienne M. Artificial intelligence and healthcare: a journey through history, present innovations, and future possibilities. Life. 2024 Apr 26;14(5):557. https://doi.org/10.3390/life14050557

[56] Wubineh BZ, Deriba FG, Woldeyohannis MM. Exploring the opportunities and challenges of implementing artificial intelligence in healthcare: A systematic literature review. InUrologic Oncology: Seminars and Original Investigations 2024 Mar 1 (Vol. 42, No. 3, pp. 48-56). Elsevier. https://doi.org/10.1016/j.urolonc.2023.11.019