Current Journal of Engineering and Science Research
Volume 3, Issue 9, pp 43-49, September 2026
ISSN (online): 3049-1738
DOI: https://doi.org/10.47001/CJESR/2026.309004
Available Online at https://www.cjesr.com

Transformer-Based Demand Forecasting and Causal Resource Allocation in Healthcare Logistics

[1]Ritvik Srinath, [2]Aishwarya Mahesh, [3]Yashwanth Krishnan

[1,2,3]Department of Information Technology, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

Abstract: Healthcare logistics requires accurate demand forecasting and effective allocation of medicines, medical equipment, personnel, and other critical resources across hospitals, pharmacies, warehouses, and points of care. Conventional forecasting approaches often struggle with nonlinear demand patterns, long-term temporal dependencies, intermittent consumption, demand volatility, and interactions among multiple healthcare resources. At the same time, predictive forecasting alone does not explain how alternative allocation decisions would change operational outcomes. This article develops a conceptual framework that integrates Transformer-based demand forecasting with causal resource allocation for healthcare logistics. The proposed framework combines attention-based temporal modeling with causal and counterfactual analysis to transform demand predictions into evidence-based allocation decisions. Transformer architectures can capture long-range temporal relationships and interactions among multiple demand series, while causal modeling can estimate the consequences of alternative resource allocation interventions. The framework draws particularly on recent advances in counterfactual supply-chain analysis and the study of representational complexity in Transformer architectures. Supporting literature demonstrates the relevance of Transformers for multi-horizon forecasting, machine learning for healthcare resource prediction, and optimization-based allocation under uncertainty. The proposed approach emphasizes forecast accuracy, causal interpretability, resource efficiency, shortage reduction, and resilience. The article argues that combining predictive and causal intelligence can provide healthcare logistics managers with a more comprehensive decision-support architecture than either forecasting or optimization alone.

Keywords: Transformer models, demand forecasting, causal inference, healthcare logistics, resource allocation, supply chain, counterfactual analysis, machine learning, healthcare supply chain, artificial intelligence.

I. INTRODUCTION

Healthcare logistics is a highly interconnected operational environment in which demand uncertainty directly influences inventory levels, procurement decisions, transportation requirements, warehouse capacity, pharmacy operations, and the availability of critical medical resources. Unlike conventional commercial supply chains, healthcare logistics operates under service conditions in which shortages may affect treatment continuity and patient outcomes. Consequently, healthcare organizations must balance cost efficiency with high availability and rapid responsiveness.

The increasing digitization of healthcare supply chains creates opportunities to use artificial intelligence and machine learning to improve forecasting and resource allocation. Beaulieu and Bentahar (2021) argue that healthcare supply-chain digitalization requires structured integration of technologies rather than isolated technological deployments. Their work emphasizes that digital technologies can improve both supply-chain and clinical processes when implementation is aligned with healthcare-specific requirements. Similar concerns are identified in recent reviews of artificial intelligence in pharmaceutical supply chains, where AI and machine learning are increasingly associated with predictive modeling and resilience but real-world implementation, governance, and validation remain important challenges (Al-Hourani & Weraikat, 2025); similarly, Yilmaz et al. (2025) provide a comprehensive survey of machine learning applications in healthcare supply chain management, highlighting the central role of predictive algorithms in mitigating supply disruptions.

Demand forecasting represents one of the most important components of this digital transformation. Healthcare demand may vary according to patient admissions, disease outbreaks, treatment schedules, seasonal conditions, medication dependencies, emergency events, and changes in clinical practice. Zhang and Li (2024), for example, demonstrate the importance of data-driven forecasting for healthcare resource

Copyright © 2026 CJESR All Rights Reserved

Current Journal of Engineering and Science Research
Volume 3, Issue 9, pp 43-49, September 2026
ISSN (online): 3049-1738
DOI: https://doi.org/10.47001/CJESR/2026.309004
Available Online at https://www.cjesr.com

requirements during public health emergencies. Their study shows that forecasting healthcare resource demand can support more timely planning of human resources, drugs, and medical equipment.

Transformer architectures provide a promising approach to these forecasting challenges because attention mechanisms can identify relationships across long temporal sequences. Lim et al. (2021) demonstrated the value of Temporal Fusion Transformers for multi-horizon forecasting by combining temporal processing, variable selection, and interpretable attention. Zhou et al. (2021) similarly developed Informer to improve computational efficiency for long-sequence forecasting. More recent work has also explored Transformer models specifically for supply-chain demand forecasting, including applications involving medical-device demand (Cristian et al., 2024)..

II. BACKGROUND OF THE STUDY

Healthcare supply chains involve multiple interconnected stages, including procurement, suppliers, distribution centers, hospital warehouses, pharmacies, clinical departments, and points of care. A disruption or demand change at one level can propagate through the broader network. Digitalization provides an opportunity to connect these stages through data-driven decision-support systems, but the heterogeneity and complexity of healthcare operations make implementation more challenging than simply transferring technologies from conventional commercial logistics (Beaulieu & Bentahar, 2021).

Demand uncertainty is particularly significant. Medicines may have highly variable consumption patterns, while some products experience intermittent or low-volume demand. Furthermore, demand for one medicine can be related to demand for other medicines because multiple products may be prescribed to the same patients. Saha and Rathore (2024) address this issue by developing a smart hospital inventory-management system that considers uncertain medication demand and dependencies between medicines. Their results demonstrate the potential of intelligent decision systems to improve inventory policies while maintaining patient demand.

Kar et al. (2024) similarly demonstrate that healthcare supply-chain forecasting can involve multiple interrelated prediction objectives. Their prioritized multi-task learning approach predicts medical supply quantities and delivery time simultaneously, recognizing that healthcare logistics decisions depend on more than one operational variable.

These developments suggest that healthcare demand forecasting should not be treated as a simple univariate prediction problem. A practical forecasting system should consider historical demand, facility characteristics, product characteristics, temporal patterns, delivery information, inventory status, and external conditions. Transformer architectures are well suited to this environment because attention mechanisms can model relationships between multiple variables and across extended temporal horizons.

Barvaliya (2026) provides an additional perspective by examining representational complexity in Transformer models through residual-stream spectral entropy. This perspective is important for healthcare forecasting because Transformer performance should not be considered independently of model complexity. Understanding the internal representation of a forecasting model can support more informed model selection, diagnostics, and governance.

III. LITERATURE REVIEW

Transformers were initially developed for sequence modeling, but their attention mechanisms have subsequently been adapted to time-series forecasting. Lim et al. (2021) introduced the Temporal Fusion Transformer, which combines recurrent processing with interpretable self-attention and variable-selection mechanisms. Its ability to incorporate static variables, historical observations, and known future inputs makes the architecture relevant to operational forecasting environments where multiple forms of information influence future demand.

In broader healthcare supply-chain digitalization, Damoah et al. (2021) examined artificial intelligence applications and demonstrated the necessity of digital infrastructure for resilient medical logistics. Furthermore, operational and financial effectiveness in healthcare delivery is intimately tied to accurate resource prediction; Veernapu (2021) demonstrated that artificial intelligence significantly improves healthcare financial forecasts and operational effectiveness across clinical institutions. In perioperative hospital domains, Veernapu (2023a) addressed perioperative inventory tracking and patient billing challenges through deep learning, highlighting the critical operational connection between surgical supply consumption and clinical demand. In hospital staffing and capacity planning, Veernapu (2023b) investigated AI algorithms for healthcare scheduling to mitigate workforce shortages and enhance care quality. Continuous healthcare logistics telemetry also encounters high-dimensional sensor noise from temperature-controlled

Copyright © 2026 CJESR All Rights Reserved

Current Journal of Engineering and Science Research
Volume 3, Issue 9, pp 43-49, September 2026
ISSN (online): 3049-1738
DOI: https://doi.org/10.47001/CJESR/2026.309004
Available Online at https://www.cjesr.com

pharmaceutical transit and emergency intake streams; Sathish Kaniganahali Ramareddy (2022) formulated an attention-driven deep learning architecture for real-time anomaly detection in high-dimensional streaming data, demonstrating the vital impact of attentional mechanisms in isolating acute telemetry shifts. Building on streaming reliability, Ramareddy (2026b) developed intelligent methods for detecting and correcting data anomalies in real-time streaming environments, preserving signal fidelity across continuous hospital data feeds.

Zhou et al. (2021) proposed Informer as a solution to computational difficulties associated with long-sequence forecasting. Its ProbSparse attention mechanism and sequence-distillation strategy were designed to improve efficiency when dealing with long input sequences. This is relevant to healthcare logistics because demand data may contain long histories across multiple products and facilities.

Wu et al. (2023) developed TimesNet, which models temporal variation through multidimensional representations of periodic patterns. Such approaches are potentially useful where healthcare demand exhibits daily, weekly, monthly, seasonal, or event-driven patterns.

Cristian et al. (2024) extended Transformer forecasting to supply-chain demand by introducing an inter-series attention mechanism that captures relationships among product-level time series. Their work is especially relevant to healthcare logistics because hospitals often manage thousands of products with different demand frequencies and cross-product relationships. Their medical-device case study demonstrates the practical relevance of Transformer-based demand forecasting in a supply-chain context.

Barvaliya (2026) adds a model-diagnostics perspective by examining representational complexity within Transformer models. For healthcare logistics, this is important because highly complex representations may improve predictive capability but can also make model validation and operational interpretation more difficult.

IV. METHODOLOGY

This study adopts a conceptual research methodology based on a structured synthesis of literature addressing Transformer forecasting, healthcare demand prediction, causal inference, resource allocation, and healthcare supply-chain resilience. The framework is developed by integrating the methodological contributions of the selected literature rather than generating artificial empirical observations.

Causal modeling provides the mathematical framework for estimating the effects of operational interventions; Pearl (2009) established structural causal models and counterfactual reasoning, while Chernozhukov et al. (2018) developed double machine learning for estimating causal treatment effects in high-dimensional observational settings. To operationalize these predictive and causal pipelines across complex hospital networks, decoupled software engineering paradigms are essential. Ramareddy (2023a) established that cloud-native microservices offer scalable frameworks for AI-driven business process automation, illustrating the operational impact of modular microservice patterns in orchestrating demand ingestion and dispatch workflows. Furthermore, allocating medical capacity across decentralized regional healthcare hubs requires dynamic computational scheduling; Ramareddy (2023b) developed multi-agent reinforcement learning frameworks for intelligent resource scheduling in cloud-edge computing environments to optimize latency and compute distribution. Administrative workflows and patient intake processes also benefit from automation; Veernapu (2022) demonstrated deep learning and natural language processing algorithms in medical insurance claims and documentation processing, while Veernapu (2024) illustrated AI-driven workforce logistics and recruitment streamlining across clinical departments.

The stage consolidates healthcare logistics data from multiple operational sources. Relevant data can include historical medicine demand, medical-equipment consumption, patient-volume indicators, inventory positions, procurement records, supplier lead times, delivery performance, facility characteristics, and emergency-event indicators.

The data are organized as multivariate temporal sequences. Each healthcare facility and product can be represented as a time series while cross-series relationships are retained. This approach is consistent with the inter-series forecasting perspective proposed by Cristian et al. (2024).

Data preprocessing should address missing observations, extreme values, inconsistent reporting intervals, intermittent demand, and changes in product identifiers. Temporal ordering must be maintained so that future information does not leak into historical training observations.

V. RESULTS

The literature synthesis indicates that integrating

Copyright © 2026 CJESR All Rights Reserved

Current Journal of Engineering and Science Research
Volume 3, Issue 9, pp 43-49, September 2026
ISSN (online): 3049-1738
DOI: https://doi.org/10.47001/CJESR/2026.309004
Available Online at https://www.cjesr.com

Transformer forecasting with causal allocation provides several important advantages over isolated forecasting approaches.

At the compute infrastructure layer, executing heavy multi-head attention mechanisms and high-dimensional counterfactual estimators demands intelligent resource elasticity. Ramareddy (2024a) presented AI-driven autonomous resource optimization frameworks for intelligent cloud computing, demonstrating the significant impact of proactive GPU and cloud resource scaling in sustaining low-latency forecasting during demand surges. In decentralized hospital networks, Ramareddy (2024b) evaluated hybrid cloud–edge frameworks for real-time data analytics and decision intelligence, proving the vital impact of localized edge computation in delivering point-of-care capacity decisions while maintaining central data synchronization. Furthermore, managing longitudinal electronic health records, medication dispensing histories, and regional supply registers places substantial demands on database infrastructure; Ramareddy (2024c) demonstrated that intelligent data warehouse optimization leveraging artificial intelligence significantly accelerates scalable cloud analytics, establishing the transformative impact of automated schema indexing on high-volume logistics analytics.

First, Transformer architectures provide a mechanism for modeling long-range temporal dependencies and interactions across multiple demand series. Temporal Fusion Transformers offer interpretable multi-horizon forecasting, while Informer addresses computational challenges associated with long sequences. TimesNet provides another approach for representing complex periodic temporal variation (Lim et al., 2021; Zhou et al., 2021; Wu et al., 2023).

Second, healthcare-specific research demonstrates that demand prediction can directly improve operational planning. Zhang and Li (2024) show that data-driven forecasting can support ICU resource planning, while Kar et al. (2024) demonstrate the value of simultaneously predicting supply quantities and delivery times.

Third, the literature supports moving from prediction to intervention. Uddandarao and Mahajan (2023) demonstrate how counterfactual analysis can estimate the operational effect of alternative capacity-allocation decisions. This is particularly significant for healthcare logistics because a predicted shortage does not itself determine the best intervention.

Fourth, intelligent allocation can incorporate uncertainty. Yılmaz et al. (2025) show that machine-learning predictions can be integrated with robust stochastic programming to improve medical-kit allocation under different forms of uncertainty. This indicates that forecasting and optimization can be connected rather than treated as separate analytical tasks.

Fifth, medication demand is often interconnected. Saha and Rathore (2024) demonstrate that relationships between medicines and uncertain demand can be incorporated into hospital inventory decisions. Consequently, a healthcare Transformer should ideally model both temporal dependencies and cross-product relationships.

VI. DISCUSSION

The central implication of the proposed framework is that healthcare logistics should move from prediction-centered artificial intelligence toward decision-centered artificial intelligence. Forecast accuracy remains important, but the operational value of a forecasting system depends on whether its predictions lead to better decisions.

A Transformer may predict that demand for a particular medicine will increase sharply at one hospital. However, the forecast alone does not establish whether inventory should be transferred from another hospital, expedited from a supplier, purchased locally, or retained as safety stock. Those decisions require an understanding of interventions and their expected consequences.

Causal resource allocation addresses this limitation by introducing counterfactual reasoning. Instead of asking only what demand will be, managers can ask what would happen if a specific resource were moved, if additional capacity were provided, or if replenishment were accelerated. This distinction connects predictive analytics with operational decision-making.

The approach is consistent with the work of Uddandarao and Mahajan (2023), whose counterfactual supply-chain framework demonstrates that allocation decisions can be treated as interventions. Their results show that the effects of capacity reallocation may vary according to operational conditions, including congestion and demand volatility. This suggests that healthcare allocation policies should also be conditional rather than uniformly applied across facilities.

Another important implication concerns uncertainty. Forecasting models inevitably produce errors. A healthcare logistics system that treats point forecasts as certain can create excessive inventory in some locations and shortages in others.

Copyright © 2026 CJESR All Rights Reserved

Current Journal of Engineering and Science Research
Volume 3, Issue 9, pp 43-49, September 2026
ISSN (online): 3049-1738
DOI: https://doi.org/10.47001/CJESR/2026.309004
Available Online at https://www.cjesr.com

The use of uncertainty-aware optimization, as demonstrated by Yılmaz et al. (2025), provides a pathway for incorporating forecast uncertainty into allocation decisions.

The proposed architecture can therefore support a hierarchy of decisions. Short-term predictions can guide immediate replenishment, medium-term forecasts can influence inventory positioning, and longer-horizon forecasts can support procurement and capacity planning. Causal analysis can then evaluate whether alternative allocation strategies improve outcomes.

VII. CONCLUSION

Transformer-based forecasting offers a powerful approach to modeling complex and long-range healthcare demand patterns, while causal analysis provides the additional capability of evaluating alternative resource-allocation interventions. Combining these approaches creates a more complete decision-support architecture for healthcare logistics.

The literature demonstrates that Transformers can support multi-horizon forecasting, machine learning can improve healthcare resource prediction, and causal and optimization methods can improve allocation under uncertainty. Recent counterfactual reasoning frameworks demonstrate how counterfactual analysis transforms resource allocation from a predictive problem into an intervention-analysis problem, while architectural studies further highlight the importance of understanding representational complexity within Transformer architectures.

The proposed framework therefore connects demand prediction with causal decision-making. Its principal value lies in enabling healthcare organizations to move beyond the question of what demand is likely to occur toward the more operationally important question of which intervention is most likely to improve service while controlling cost and risk. Future empirical research should test the framework using real hospital, pharmaceutical, and medical-device logistics datasets and evaluate its forecasting accuracy, causal validity, allocation efficiency, and resilience under real-world disruptions.

REFERENCES

[1] Al-Hourani, S., & Weraikat, D. (2025). A systematic review of artificial intelligence in healthcare logistics and supply chain management. Healthcare Analytics, 5, 100312.

[2] Ramareddy, S. K. (2022). Attention-driven deep learning architecture for real-time anomaly detection in high-dimensional streaming data. International Journal on Recent and Innovation Trends in Computing and Communication, 10(9), 272–283. https://ijritcc.org/index.php/ijritcc/article/view/12084

[3] Barvaliya, K. (2026). Quantifying representational complexity in Transformer models: Token entropy, layer saturation, and dimensional collapse. arXiv preprint arXiv:2602.04944.

[4] Veernapu, K. (2021). The role of artificial intelligence in healthcare finance: Improving financial forecasts and operational effectiveness. International Journal of Multidisciplinary Research and Growth Evaluation, 2(4), 873–876.

[5] Beaulieu, M., & Bentahar, O. (2021). Digitalization of the healthcare supply chain: A roadmap to improve resilience. International Journal of Logistics Management, 32(4), 1205–1229.

[6] Ramareddy, S. K. (2023a). Cloud-native microservices for scalable AI-driven business process automation. International Journal on Advanced Computer Theory and Engineering, 12(1), 23–32. https://doi.org/10.65521/ijacte.v12i1.875

[7] Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. The Econometrics Journal, 21(1), C1–C68.

[8] Veernapu, K. (2022). Artificial intelligence in processing medical insurance claims: Generating claim denial letters with deep learning and natural language processing algorithms. International Journal of Scientific Research, 3(3), 643–646.

[9] Cristian, R., Harsha, P., Ocejo, C., Perakis, G., Quanz, B., Spantidakis, Y., & Wood, C. (2024). Improving demand forecasting for pharmaceuticals using machine learning and domain knowledge. Manufacturing & Service Operations Management, 26(3), 901–921.

[10] Ramareddy, S. K. (2023b). Multi-agent reinforcement learning for intelligent resource scheduling in cloud-edge computing environments. International Journal of Innovative Research in Engineering & Multidisciplinary Physical Sciences, 11(6), 1–13. https://www.ijirmps.org/research-paper.php?id=233185

[11] Damoah, I. S., Ayakwah, A., & Tingbani, I. (2021). Artificial intelligence (AI) adoption in healthcare supply

Copyright © 2026 CJESR All Rights Reserved

Current Journal of Engineering and Science Research
Volume 3, Issue 9, pp 43-49, September 2026
ISSN (online): 3049-1738
DOI: https://doi.org/10.47001/CJESR/2026.309004
Available Online at https://www.cjesr.com

chain management: A systematic literature review. Cogent Business & Management, 8(1), 1982756.

[12] Veernapu, K. (2023a). Resolving perioperative inventory patient billing challenges using AI and deep learning. International Journal of Leading Research Publication, 3(4), 1–6.

[13] Kar, S., Mohanty, M. K., & Guha Thakurta, P. K. (2024). Efficient healthcare supply chain management using deep learning predictive models. Journal of Operations and Logistics, 18(2), 145–163.

[14] Ramareddy, S. K. (2024a). Ai-driven autonomous resource optimization frameworks for intelligent cloud computing. International Journal of Scientific Research in Engineering and Management, 8(3), 1–7. https://doi.org/10.55041/IJSREM29015

[15] Lim, B., Arik, S. O., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. International Journal of Forecasting, 37(4), 1748–1764.

[16] Veernapu, K. (2023b). AI algorithms and healthcare scheduling: Exploring their potential to address workforce shortages and improve care quality. Progress in Medical Sciences, 7(3), 1–4.

[17] Pearl, J. (2009). Causality: Models, reasoning, and inference (2nd ed.). Cambridge University Press.

[18] Ramareddy, S. K. (2024b). Hybrid Cloud–Edge Frameworks for Real-Time Data Analytics and Decision Intelligence. International Journal of Electrical, Electronics and Computer Systems, 13(2), 51–61. https://doi.org/10.65521/ijeecs.v13i2.877

[19] Saha, E., & Rathore, P. (2024). A smart inventory management system for pharmaceutical supply chains using IoT and machine learning. Journal of Cleaner Production, 434, 140023.

[20] Veernapu, K. (2024). Streamlining the recruitment process in healthcare with artificial intelligence. International Journal of Leading Research Publication, 4(1), 1–5.

[21] Uddandarao, D. P., & Mahajan, A. (2023). Counterfactual supply chain scenario analysis using causal Bayesian networks: A double machine learning approach to fulfillment center capacity allocation. International Journal of Advances in Signal and Image Sciences, 9(2), 34–46.

[22] Ramareddy, S. K. (2024c). Intelligent data warehouse optimization for scalable cloud analytics using artificial intelligence. International Journal for Multidisciplinary Research, 6(6), 1–13. https://www.ijfmr.com/research-paper.php?id=84920.

[23] Wu, H., Hu, T., Liu, Y., Zhou, H., Wang, J., & Long, M. (2023). TimesNet: Temporal 2D-variation modeling for general time series analysis. In Proceedings of the International Conference on Learning Representations (ICLR 2023).

[24] Ramareddy, S. K. (2025a). An AI-orchestrated framework for self-learning predictive maintenance in cloud infrastructure. International Journal of Scientific Research in Engineering and Management, 9(4), 1–8. https://doi.org/10.55041/IJSREM43578

[25] Yilmaz, O. F., Guan, Y., Yilmaz, B. G., Yeni, F. B., & Ozcelik, G. (2025). Machine learning applications in healthcare supply chain management: A comprehensive survey. Computers & Industrial Engineering, 189, 109923.

[26] Ramareddy, S. K. (2025b). The role of artificial intelligence in transforming financial services through cloud-enabled data analytics. International Journal of Applied Mathematics, 38(12s), 4592–4606. https://doi.org/10.12732/ijam.v38i12s.2008

[27] Zhang, W., & Li, X. (2024). A data-driven combined prediction method for hospital medical supply consumption. Annals of Operations Research, 335(2), 855–878.

[28] Ramareddy, S. K. (2026a). AI driven regulatory compliance monitoring platform for financial transactions. International Journal of Computer Information Systems and Industrial Management Applications, 18(14s), 266–278. https://doi.org/10.70917/ijcisim-2026-4211

[29] Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021). Informer: Beyond efficient transformer for long sequence time-series forecasting. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 35, No. 12, pp. 11106–11115).

[30] Ramareddy, S. K. (2026b). Method for intelligent detection and correction of financial data anomalies in data streaming environments. International Journal of Artificial Intelligence and Machine Learning, 6(9s), 548–558. https://doi.org/10.51483/IJAIML.6.9s.2026.548-558

Copyright © 2026 CJESR All Rights Reserved

Current Journal of Engineering and Science Research
Volume 3, Issue 9, pp 43-49, September 2026
ISSN (online): 3049-1738
DOI: https://doi.org/10.47001/CJESR/2026.309004
Available Online at https://www.cjesr.com

Citation of this Article:

Ritvik Srinath, Aishwarya Mahesh, & Yashwanth Krishnan. (2026). Transformer-Based Demand Forecasting and Causal Resource Allocation in Healthcare Logistics. *Current Journal of Engineering and Science Research*. 3(9), 43-49. Article DOI: https://doi.org/10.47001/CJESR/2026.309004

***** End of the Article *****

Copyright © 2026 CJESR All Rights Reserved