
EARLY PREDICTION OF SEPSIS: FEATURE-ALIGNED TRANSFER LEARNING

Oyindolapo O. Komolafe

School of Physical Therapy,
Faculty of Health Sciences,
Western University, London, ON, CA
okomola@uwo.ca

Zhimin Mei

Department of Computer Science,
Western University
London, ON, CA
zmei22@uwo.ca

David Morales Zarate

Health and Rehabilitation Sciences
Faculty of Health Sciences,
Western University, London, ON, CA
omorales@uwo.ca

Gregory William Spangenberg

Mechanical Engineering
Western University
London, ON, CA
gspangen@uwo.ca

May 7, 2025

ABSTRACT

Sepsis is a life-threatening medical condition that occurs when the body has an extreme response to infection, leading to widespread inflammation, organ failure, and potentially death. Because sepsis can worsen rapidly—sometimes within hours—early detection is critical to saving lives. However, current diagnostic methods often identify sepsis only after significant damage has already occurred. Our project aims to address this challenge by developing a machine learning-based system to predict sepsis in its early stages, giving healthcare providers more time to intervene.

A major problem with existing models is the wide variability in the patient information—or features—they use, such as heart rate, temperature, and lab results. This inconsistency makes models difficult to compare and limits their ability to work across different hospitals and settings. To solve this, we propose a method called **Feature-Aligned Transfer Learning (FATL)**, which identifies and focuses on the most important and commonly reported features across multiple studies, ensuring the model remains consistent and clinically relevant.

Moreover, most existing models are trained on narrow patient groups, leading to population bias. FATL addresses this by combining knowledge from models trained on diverse populations, using a weighted approach that reflects each model’s contribution. This makes the system more generalizable and effective across different patient demographics and clinical environments. FATL offers a practical and scalable solution for early sepsis detection, particularly in hospitals with limited resources, and has the potential to improve patient outcomes, reduce healthcare costs, and support more equitable healthcare delivery.

Keywords sepsis · early prediction · weight transfer · transfer learning

1 Introduction

Sepsis is a life-threatening condition that occurs when the body’s immune system overreacts to an infection [1]. Under normal circumstances, the body releases specific chemicals and proteins to regulate immune responses, control inflammation, and promote healing. However, in sepsis, this process becomes dysregulated, triggering an excessive and widespread inflammatory response throughout the body. Figure 1 shows the damage that sepsis can cause to different organs and tissues in the human body, reflecting its dangerousness.

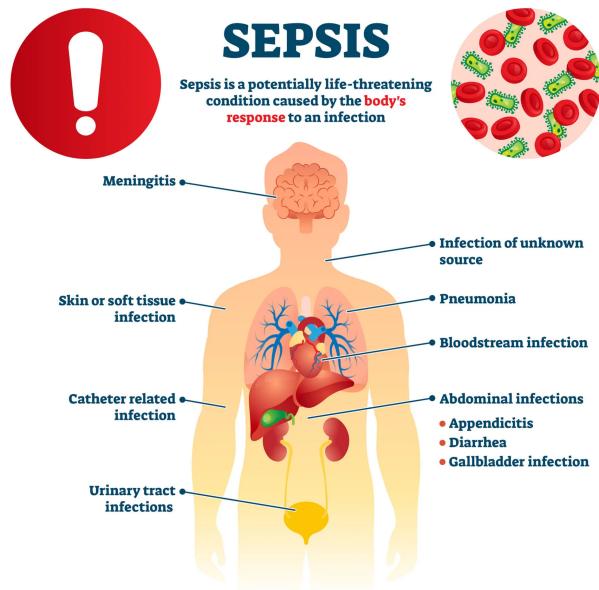


Figure 1: Introduction of sepsis

The health risks associated with sepsis are severe and far-reaching. Often referred to as blood poisoning, sepsis can disrupt blood flow to vital organs and cause abnormal clotting, which may impair organ function. While early detection and treatment can often lead to recovery, untreated or advanced sepsis can progress rapidly to irreversible multiple organ failure, septic shock—the most critical phase—and ultimately, death.

1.1 Societal and Clinical Importance of Sepsis

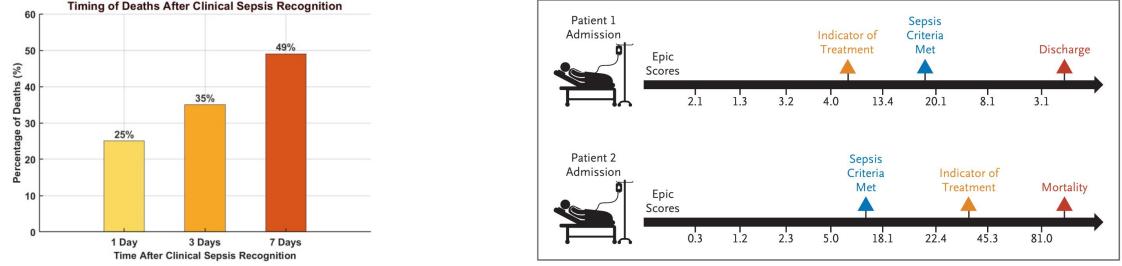
Sepsis is a leading cause of critical illness and death worldwide [2]. Estimates indicate 48.9 million cases of sepsis occurred in 2017, with 11 million or 20% of deaths worldwide [3]. The World Health Organization (WHO) has recognized sepsis as a global health priority and urged member states to improve its prevention, diagnosis, and management. Not only is sepsis deadly, but it is also costly. Managing sepsis often requires intensive care, prolonged hospital stays, and complex interventions. In high-income countries, the average hospital cost per sepsis case can exceed \$32,000 (USD) [4]. According to Sepsis Canada, sepsis imposes an estimated \$1.7 billion (CAD) annually in direct costs [5], with around \$1 billion per year in Ontario [6]. Beyond acute care, many sepsis survivors experience long-term consequences, which create additional societal costs for rehabilitation and support.

1.2 Importance of early prediction of sepsis

Sepsis has two critical characteristics: complex diagnosis and rapid progression. Diagnosing sepsis requires extensive testing, including blood tests, cultures, imaging, and organ function assessments such as the SOFA score [7]. These procedures help identify infection sources and evaluate the extent of organ dysfunction. In addition, sepsis progresses quickly—within 12 hours, a patient's condition can deteriorate severely. According to Masmali et al. [8], 25% of sepsis-related deaths occur within one day of diagnosis, increasing to 49% by day seven shown by Figure 2a. This highlights the urgent need for early detection and intervention.

Figure 2b illustrates two contrasting cases. Patient 1 received early treatment before meeting clinical criteria and recovered. Patient 2 was treated only after diagnosis, by which point organ failure had likely set in, leading to death. Early prediction is therefore essential for improving patient outcomes.

Therefore, leveraging rapidly advanced AI technologies to develop models that can accurately predict sepsis several hours in advance holds great promise for improving patient outcomes.



(a) The mortality rate increases rapidly with the time of clinical sepsis diagnosis.

(b) Different conditions of two patients

Figure 2: Comparison of mortality and early prediction for patients

2 Research Gaps

Despite growing interest in machine learning for sepsis prediction, persistent research gaps hinder its clinical adoption. Studies often differ in data, features, and sepsis definitions, complicating comparisons and reproducibility—only 1 of 28 studies followed standard reporting guidelines [9]. Most models are validated on retrospective North American datasets, limiting generalization to diverse or resource-limited settings [10]. Real-world deployment is rare, with high false-positive rates in EHR-based tools leading to alarm fatigue and clinician mistrust [11].

Crucial challenges remain, including the lack of randomized controlled trials, limited contextual validation in settings like emergency departments, and data quality issues such as missing or inconsistent EHR entries [11, 10]. Deep learning models face scrutiny over interpretability, with clinicians needing explainable outputs from tools like SHAP or LIME [12]. Ethical concerns about biased training data and device inaccuracies raise equity issues, while the long-term impacts of false positives are understudied [13].

Few models undergo robust external validation; widely used tools like the Epic Sepsis Model often perform poorly in new environments [14]. Advancing the field requires standardization, multi-center prospective validation, better data access, and innovations in temporal modeling and ensemble methods to ensure clinical relevance and equity [15].

3 Proposed Method & Machine Learning Solution

According to a systematic review by Fleuren et al. (2020)[9], a wide range of studies have addressed sepsis prediction, including early detection and neonatal mortality. However, these studies often rely on different features and clinical scores, such as the SOFA score, leading to inconsistency. While the SOFA score is useful for assessing organ dysfunction in ICU settings, it may not be optimal for early detection, as it typically identifies sepsis only after significant deterioration has occurred. Early intervention is critical in managing sepsis, and models that depend on overt signs of organ failure may detect the condition too late. These differences in feature selection and population focus make it difficult to generalize predictive models across datasets and clinical environments. As illustrated in Figure 3, we propose a unified approach that integrates knowledge from multiple source models to improve generalizability. Our method is based on a **feature-aligned transfer learning (FATL)** framework.

- First, we identify biological features that are commonly emphasized across prior models. These shared features act as robust indicators and can help prioritize patient testing, especially in settings with limited medical resources.
- Second, we address population bias by leveraging source models trained on diverse cohorts. Instead of relying on a single dataset, we combine model parameters using a weighted scheme to capture a broader representation of patient profiles.

By aligning features and aggregating knowledge from diverse sources, FATL enables a more reliable and transferable solution for early sepsis prediction.

3.1 Feature-Aligned Transfer Learning (FATL)

FATL (4) is a machine learning methodology that combines transfer learning with a focus on feature selection. In FATL, the process involves:

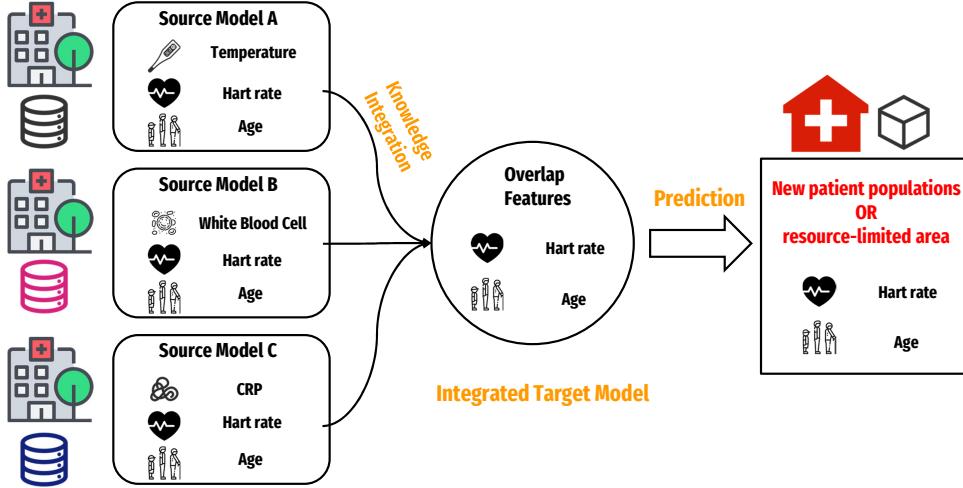


Figure 3: The architecture of our system

Conceptual Architecture Diagram of FATL

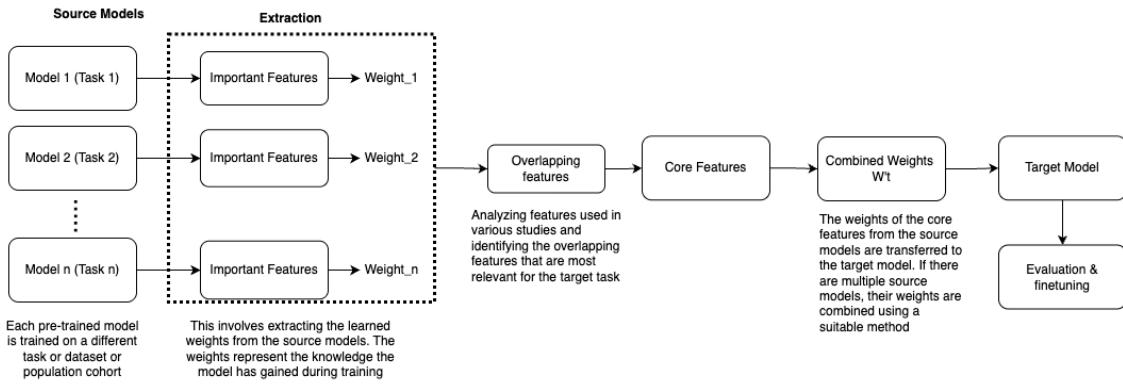


Figure 4: Conceptual Architecture Diagram of FATL

- Identifying key features that are consistently important across various studies or datasets related to the target prediction task.
- Transferring knowledge from source models to a target model, with an emphasis on aligning the transferred knowledge with these key features

The goal of FATL is to improve prediction accuracy, generalization, and robustness by focusing on the most relevant information and leveraging knowledge from diverse sources.

3.1.1 Feature Selection and Harmonization

This process involves: (1)Review of the literature of the focus area to identify the features considered important in various studies. (2)Analyze the selected studies and identify features that consistently appear as significant predictors. (3)Harmonize the overlapping features to ensure consistency in definition, units or scales across the datasets. (4)Identify and handle data quality

3.1.2 Weight Transfer

Weight transfer is a machine learning technique where a model trained on one task is reused for a new task, typically when the new task has less data. It helps the new model start with useful, general features learned from a larger dataset. In sepsis detection, weight transfer allows the combining of models trained on different features, improving sensitivity

to early signs and enhancing generalization across diverse patients and clinical settings. Weight transfer for *target model* with various tasks *source models* is defined as:

$$W't = \sum_{i=1}^n \alpha_i \cdot W_i \quad (1)$$

$W't$ is the initial weight of the target model after the transfer.

α_i weight transfer coefficient representing the importance or contribution of the i-th source model. W_i weights from the i-th source model, where (1=1,2,...n) i.e weights from the various tasks n.

The weights of each source model W_i are multiplied by a coefficient α_i that determines how much influence that model has on the initial weights of the target model.

α_i coefficient allows for flexibility in giving more or less importance to different source models based on their relevance to the target task. If $\alpha_i = \frac{1}{n}$ for all (i), it becomes a simple average of all source model weights.

Equation 1 allows weight transfer from multiple sources and fine-tuning of each source contribution.

3.1.3 FATL formula

Therefore FATL can be defined as:

$$W't = \sum_{i=1}^n \alpha_i \cdot (W_i \odot F) - \lambda \cdot (b_t - \frac{1}{n} \sum_{i=1}^n b_i) \quad (2)$$

F is a filter that allows the transfer of weights from source models focused on the most important features.

F is an array of weights representing the importance of the j-th feature, where ($j = 1, 2, \dots, m$) (total number of features).

If F_j is close to 1, the corresponding weights are transferred; if it's close to 0, they are suppressed.

α_i weight transfer coefficient representing the importance or contribution of the i-th source model.

$\lambda \cdot (b_t - \frac{1}{n} \sum_{i=1}^n b_i)$ is the bias regularization term.

λ Bias penalty coefficient, controlling the strength of the penalty. A larger λ enforces a stronger similarity between the target model's bias and the source models' average bias. b_t Bias vector of the target model. b_i Bias vector of the i-th source model.

3.2 Implementation of FATL for Sepsis Use case

To begin the process, the PubMed database was searched for publications done between 2020-2025 using the keywords: Machine Learning, Sepsis, Detection/Prediction. Finally, 18 papers were selected for research. From the reviews, the following were observed: (1) Most of the papers didn't report the weights/coefficients of the features. (2) ML models are not well documented and, therefore, not reproducible. Pre-processing done on the features is sparsely documented. (3) When using multiple ML, each model's feature importance isn't disclosed.

3.3 Preliminary Result

Based on an analysis of the 18 selected articles, we identified the top 10 most important biological characteristics, as shown in Figure 5. These features fall into three categories: vital signs, laboratory values, and demographic information. They can be considered the key indicators to help in the detection and prediction of sepsis. In resource-limited medical settings, focusing on these features can help prioritize and guide early diagnosis and intervention efforts.

3.4 Conclusion, Challenges and Limitations

The FATL (Feature-Aware Transfer Learning) methodology combines the strengths of weight transfer - leveraging knowledge from diverse data - and feature selection -focusing on the most relevant information - to develop a more accurate and robust model. FATL is applicable to a variety of problems, particularly in healthcare and scenarios where data is scarce or heterogeneous such as medical diagnosis, image recognition and Natural Language Processing. FATL method can improve accuracy, by focusing on overlapping important features, models can prioritize the most relevant

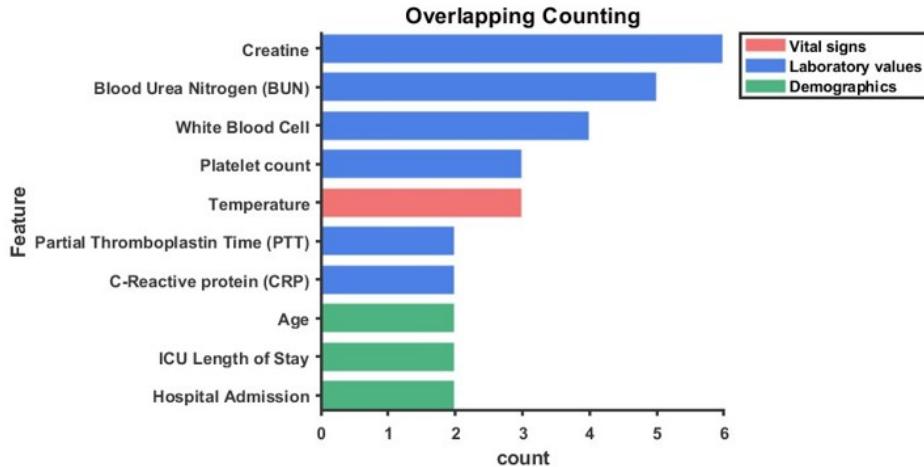


Figure 5: Overlapping Features

information, potentially leading to higher accuracy. Models can generalize better due to weight transfer from multiple source models and this can lead to faster convergence of model. Subsequently, focusing on clinically relevant and consistently identified features can increase the model's acceptance and utility in clinical practice. On the flip side, FATL relies on weights and known bias to be efficient, therefore, inconsistencies or biases in the source model can affect model performance.

3.5 GenAI tool usage

In this report, we used GPT-4o [16] to help polish sentences and correct grammatical errors, but we already wrote the content we wanted to express.

4 Stakeholder Analysis

If this project advances to clinical testing and implementation, it will require collaboration across multiple domains. Key stakeholders include frontline healthcare providers, who would use the system and provide workflow feedback, and patients and families, who would guide design decisions on transparency and privacy. In Canada, groups like Sepsis Canada involve patient partners in improving detection and care. Hospital leaders, data scientists, IT professionals, and policymakers would support safety, validation, integration, and regulatory compliance. Identifying these stakeholders early lays the groundwork for future implementation.

References

- [1] Nita Crighton. What is sepsis?, 2022. Accessed: April 9, 2025. Available at: <https://blog.encompasshealth.com/2022/09/21/what-is-sepsis/>.
- [2] Carolin Fleischmann-Struzek and Kristina Rudd. Challenges of assessing the burden of sepsis. *Medizinische Klinik-Intensivmedizin und Notfallmedizin*, 118(Suppl 2):68–74, 2023.
- [3] Kristina E Rudd, Sarah Charlotte Johnson, Kareha M Agesa, Katya Anne Shackelford, Derrick Tsoi, Daniel Rhodes Kievlanc, Danny V Colombara, Kevin S Ikuta, Niranjan Kissoon, Simon Finfer, et al. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the global burden of disease study. *The Lancet*, 395(10219):200–211, 2020.
- [4] Habibollah Arefian, Steffen Heublein, André Scherag, Frank Martin Brunkhorst, Mustafa Z Younis, Onnen Moerer, Dagmar Fischer, and Michael Hartmann. Hospital-related cost of sepsis: a systematic review. *Journal of Infection*, 74(2):107–117, 2017.
- [5] Sepsis Canada. About sepsis canada, 2024. Accessed: April 8, 2025. Available at: <https://www.sepsiscanada.ca/about/about-sepsis-canada>.

- [6] Kelly Farrah, Lauralyn McIntyre, Christopher J Doig, Robert Talarico, Monica Taljaard, Murray Krahn, Dean Fergusson, Alan J Forster, Doug Coyle, and Kednapa Thavorn. Sepsis-associated mortality, resource use, and healthcare costs: a propensity-matched cohort study. *Critical Care Medicine*, 49(2):215–227, 2021.
- [7] Rui Moreno, Andrew Rhodes, Lise Piquilloud, Glenn Hernandez, Jukka Takala, Hayley B Gershengorn, Miguel Tavares, Craig M Coopersmith, Sheila N Myatra, Mervyn Singer, et al. The sequential organ failure assessment (sofa) score: has the time come for an update? *Critical care*, 27(1):15, 2023.
- [8] O. E. Masmali, M. Alageeli, A. H. Zarban, S. A. Munhish, M. Y. Sahloli, M. H. Sanea, A. A. Moafa, H. M. Halawi, A. A. Zurayqan, and K. H. Alhazmi. Epidemiology of pediatric sepsis in the pediatric intensive care unit - a systematic review. *International Journal of Medicine in Developing Countries*, 8(11):3326–3334, 2024.
- [9] Lucas M Fleuren, Thomas LT Klausch, Charlotte L Zwager, Linda J Schoonmade, Tingjie Guo, Luca F Roggeveen, Eleonora L Swart, Armand RJ Girbes, Patrick Thoral, Ari Ercole, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intensive care medicine*, 46:383–400, 2020.
- [10] Elena Giovanna Bignami, Michele Berdini, Matteo Panizzi, Tania Domenichetti, Francesca Bezzi, Simone Allai, Tania Damiano, and Valentina Bellini. Artificial intelligence in sepsis management: An overview for clinicians. *Journal of Clinical Medicine*, 14(1):286, 2025.
- [11] Aaron Boussina, Supreeth P Shashikumar, Atul Malhotra, Robert L Owens, Robert El-Kareh, Christopher A Longhurst, Kimberly Quintero, Allison Donahue, Theodore C Chan, Shamim Nemati, et al. Impact of a deep learning sepsis prediction model on quality of care and survival. *NPJ digital medicine*, 7(1):14, 2024.
- [12] Fang Li, Shengguo Wang, Zhi Gao, Maofeng Qing, Shan Pan, Yingying Liu, and Chengchen Hu. Harnessing artificial intelligence in sepsis care: advances in early detection, personalized treatment, and real-time monitoring. *Frontiers in Medicine*, 11:1510792, 2025.
- [13] Jeffrey Smith, Andre Holder, Rishikesan Kamaleswaran, and Yao Xie. Detecting algorithmic bias in medical-ai models using trees. *arXiv preprint arXiv:2312.02959*, 2023.
- [14] Andrew Wong, Erkin Otles, John P Donnelly, Andrew Krumm, Jeffrey McCullough, Olivia DeTroyer-Cooley, Justin Pestrue, Marie Phillips, Judy Konye, Carleen Penzoza, et al. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. *JAMA internal medicine*, 181(8):1065–1070, 2021.
- [15] SherAli Bomrah, Mohy Uddin, Umashankar Upadhyay, Matthieu Komorowski, Jyoti Priya, Eshita Dhar, Shih-Chang Hsu, and Shabbir Syed-Abdul. A scoping review of machine learning for sepsis prediction-feature engineering strategies and model performance: a step towards explainability. *Critical Care*, 28(1):180, 2024.
- [16] OpenAI. Gpt-4 research index, 2024. Accessed: April 9, 2025. Available at: <https://openai.com/index/gpt-4-research/>.