Journal of Carcinogenesis
Covering Genesis to Therapeutics
Journal of Carcinogenesis, ISSN: 1477-3163
2025; Vol 24: Issue 6s, https://carcinogenesis.com/

AI-Powered Early Warning Systems for Clinical Deterioration Significantly Improve Patient Outcomes: A Meta-Analysis

Ramlah mohmmed alobaid[1], sajeda naji yousef alamer[2], zahra ali alsultan[3], Hawra Mohammed Alobaid[4], Zahra Mohammed Al-Obaid[5], Faeqah Taha Saleh Sharif[6], Yousef Ali Hamed Alshehri[7], Sarah saud alruwished[8]

[1]staff nurse, aljaber ent-oph hospital

Email: ralobaid@moh.gov.sa

[2]staff nurse, King Faisal General Hospital

Email: snalamer@moh.gov.sa

[3]staff nurse, King Faisal General Hospital

Email: zaaalsultan@moh.gov.sa

[4]Nursing specialist, Khobash General Hospital

Email: Hamalobaid@moh.gov.sa

[5]Nursing technician, Salwa General Hospital

Email: Zalobaid@Moh.gov.sa

[6]Staff Nurse, King Fahd General Hospital Jeddah

Email: fshreef@moh.gov.sa

[7]Staff Nurse, King Abdulaziz International Airport in Jeddah - Nursing Services.

Email: yalshehry@moh.gov.sa

[8]Staff nurse, King faisal general hospital

Email: Salruwished@mog.gov.sa

ABSTRACT

Background: Early observation of clinical worsening is critical for reducing morbidity and mortality in hospitalized patients. Conventional early warning scores have limited accuracy, while artificial intelligence–powered early warning systems (AI-EWS) may offer improved predictive value.

Objectives: To estimate the influence of AI-EWS on case results, involving mortality, intensive care unit (ICU) transfer, and duration of hospitalization.

Methods: This systematic review and meta-analysis have been done after PRISMA guidelines. Five investigations (2013–2024) involving 95,162 patients were included. Eligible studies compared AI-EWS with standard care or conventional scoring systems and reported mortality, ICU transfer, or length of stay. Data extraction was performed independently by 2 reviewers. Risk of bias has been evaluated utilizing the Cochrane instrument for randomized trials and the Newcastle–Ottawa Scale for observational studies. Random-influences models have been utilized for pooled analysis.

Results: AI-EWS significantly reduced all-cause mortality (OR = 0.76; ninety-five percent confidence interval: 0.63–0.91; p equal to 0.004). An insignificant variance has been found for ICU transfers (OR = 0.90; ninety-five percent confidence interval: 0.76–1.07; p equal to 0.22). Duration of stay in the hospital was modestly reduced in AI-EWS groups (MD = –0.35 days; ninety-five percent confidence interval: –0.68 to –0.01; p = 0.04). Risk of bias was low to moderate, mainly due to heterogeneity in study design.

Conclusion: AI-EWS are associated with lower mortality and shorter hospital stays compared with conventional systems, though their effect on ICU transfers remains uncertain. Larger high-quality trials are required to confirm these findings.

Keywords: *Artificial intelligence, early warning system, clinical deterioration, mortality*.

How to Cite: Ramlah mohmmed alobaid, sajeda naji yousef alamer, zahra ali alsultan, Hawra Mohammed Alobaid, Zahra Mohammed Al-Obaid, Faeqah Taha Saleh Sharif, Yousef Ali Hamed Alshehri, Sarah saud alruwished., (2025) AI-Powered Early Warning Systems for Clinical Deterioration Significantly Improve Patient Outcomes: A Meta-Analysis, *Journal of Carcinogenesis*, *Vol.24, No.6s*, 118-124.

1. INTRODUCTION

Physiological signs frequently exhibit alterations prior to clinical deterioration. Inability to promptly recognize these alterations may lead to unfavorable outcomes. The growing focus on AI for forecasting clinical deterioration in hospitalized cases is significant. Nonetheless, the efficacy of AI-based predictive models & their capacity for improving case prognosis is still uncertain **(1)**.

Presently, initiatives to identify early clinical deterioration in hospitalized cases rely on artificially computed ratings, which are sluggish, ineffective, and demand substantial investment. Traditional scoring systems, including NEWS, EWS, APACHE II, and APACHE III, often utilize characteristics such as the Glasgow Coma Scale (GCS), age, & critical signs for assessment upon admission, through alterations in medical status, & at other designated intervals **(2)**. Empirical facts indicate that these routine scoring systems can partially detect the deterioration of a case's condition. Nevertheless, the restricted information dimensions & the overgeneralization of clinical scenarios characteristic of traditional scoring systems may neglect individual variances & the intricacies of clinical settings, hence exhibiting specific limits **(3)**.

Currently, there is an increasing interest in utilizing AI for the early detection of clinically worsening cases. Machine learning algorithms, involving support vector machines, logistic regression, and neural networks, have demonstrated efficacy in certain retrospective research utilizing databases. They exhibit elevated specificity, accuracy, and sensitivity in forecasting hospitalization length and seven-day/twenty-eight-day death rates **(4)**.

In theory, AI-based models demonstrate superior adaptability and yield more precise expectations compared to traditional scoring systems. They provide real-time continuous monitoring, permitting the prompt identification of alterations in a case's state **(5)**.

AI-based models can more precisely detect clinical worsening, facilitating prompt interventions to enhance patient results. Simultaneously, enhanced sensitivity and specificity might diminish false positives, resulting in a more judicious distribution of scarce medical resources. Consequently, precisely evaluating the reliability of AI-based models & their performance disparities relative to conventional scoring systems is essential for validating the feasibility of AI-based models **(6&7)**.

This research aimed to assess the influence of AI-EWS on patient outcomes, involving death, transfer to the ICU, and duration of hospital stay.

2. PATIENTS AND METHODS

This meta-analysis was done in line with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. A comprehensive literature search has been made across Web of Science databases, Cochrane Library, Embase, and PubMed to recognize researches published from 2013 to 2024 that evaluated AI-EWS for predicting clinical deterioration in hospitalized patients. Search terms included combinations of artificial intelligence, early warning systems, machine learning, clinical deterioration, mortality, ICU transfer, and length of stay.

3. ELIGIBILITY CRITERIA:

Inclusion criteria: Studies were included if they involved adult inpatients, compared AI- or machine learning–based early warning models with traditional scoring systems or standard care, and reported at least one relevant outcome: mortality, ICU transfer, or hospital length of stay. Randomized controlled cohort studies, trials, and pre-post intervention investigations have been considered. **Exclusion criteria**: Studies have been excluded if they lacked comparative data, included pediatric populations, or were reviews, case reports, or editorials.

Data extraction and quality assessment: Two independent reviewers extracted information on research design, population characteristics, type of AI model, comparator, and clinical outcomes. Discrepancies were resolved by consensus. **Risk of bias assessment**: The methodological quality of the involved investigations was estimated utilizing proven methodologies suitable for the research design. Randomized controlled trials have been estimated with the Cochrane Risk of Bias tool, examining areas involving missing outcome data, randomization, allocation concealment, blinding, and selective reporting. Observational studies have been estimated utilizing the Newcastle–Ottawa Scale (NOS), which assigns ratings based on participant selection, cohort comparability, and outcome measurement. Studies have been categorized as having moderate, low, or great risk of bias. Disputes have been settled by dialogue with a 3rd evaluator.

Outcomes: The 1^{ry} finding was all-cause in-hospital mortality. 2^{ry} results involved intensive care unit transfer and hospital length of stay.

Statistical analysis

All analysis of data has been conducted utilizing Review Manager version 5.4.1. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014. We computed the odds ratio with a ninety-five percent CI for binary outcomes. We computed the mean difference with a ninety-five percent CI for continuous outcomes. To determine the overall influence and estimate the 95% CI, we employed a fixed-influence model utilizing the Mantel-Haenszel method in the lack of heterogeneity among studies. A random-influences model utilizing the DerSimonian and Laird approach was selected. The heterogeneity among studies has been assessed utilizing the Q statistic and I² test, which quantify the percentage of variability in the influence estimates. A P below 0.05 has been deemed significant.

4. RESULTS

5 trials were chosen for the current analysis, encompassing a total of 95,162 cases. The publication years spanned from 2013 to 2024. Four studies have been undertaken in the USA. The baseline characteristics of the investigations included are illustrated in **Table 1**.

Author, year	**year**	**country**	**Study period**		**Study design**	**Sample Size**		
			from	**to**		**MLM group**	**No-MLM group**	**total**
Levin, (8)	2024	USA	2019	2020	Nonrandomized Clustered Pragmatic Clinical Trial	1488	1252	2740
Bassin, (9)	2023	Australia	2019	2021	Single hospital, pre-post study	14540	14,099	28639
Winslow, (10)	2022	USA	2016	2018	Multicenter clinical intervention trial	3191	3490	6681
Escobar, (11)	2020	USA	2016	2019	Large, multicenter cohort	13274	23797	37071
Bailey, (12)	2013	USA	2007	2011	Randomized, controlled crossover study	9911	10120	20031

Table2. Patient's characteristics:

The average age in the investigated groups was 62.3 years, with a vary from twenty-five to seventy-six. Gender was recorded in two investigations, comprising 29,604 males and 27,498 females, as seen in **Table 2**.

Author, year	**age (years)**						**Sex**					
	MLM group			**No-MLM group**			**MLM group**			**No-MLM group**		
	mean	**SD**	**total**	**mean**	**SD**	**total**	**Male**	**Female**	**total**	**Male**	**Female**	**total**
Escobar, (11)	68.4	15.6	13274	67.2	16.2	23797	7314	5960	13274	12922	10875	23797
Bailey, (12)	57	19.3	9911	56.7	18.5	10120	4603	5308	9911	4765	5355	10120

Mortality:

Five studies stated (mortality), and all may be applied. A significant heterogeneity was recognized. Thus, a random-influence model was applied for analysis ($I^2 = 85\%$, p below 0.0001). The combined odds ratio and 95% CI were 0.76 (0.63 to 0.91). The combined finding reveals a statistically significant difference between groups according to mortality (Z = 2.89, P = 0.004).

Study or Subgroup	MLM group Events	MLM group Total	No-MLM group Events	No-MLM group Total	Weight	Odds Ratio M-H, Random, 95% CI
Bailey, 2013	223	9911	227	10120	20.4%	1.00 [0.83, 1.21]
Bassin, 2023	133	14540	151	14099	18.3%	0.85 [0.67, 1.08]
Escobar, 2020	1301	13274	3420	23797	24.7%	0.65 [0.61, 0.69]
Levin, 2024	81	1488	83	1252	14.9%	0.81 [0.59, 1.11]
Winslow, 2022	281	3191	485	3490	21.7%	0.60 [0.51, 0.70]
Total (95% CI)		**42404**		**52758**	**100.0%**	**0.76 [0.63, 0.91]**
Total events	2019		4366			

Heterogeneity: Tau² = 0.04; Chi² = 26.02, df = 4 (P < 0.0001); I² = 85%
Test for overall effect: Z = 2.89 (P = 0.004)

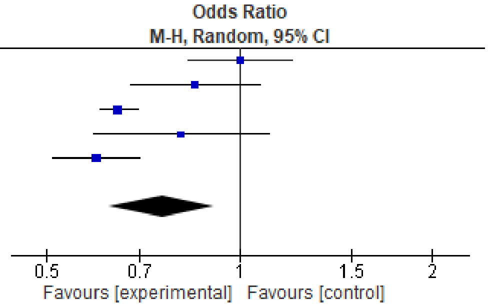

Figure (1): Forest plot of Mortality reveals a statistically significant distinction between MLM group and No-MLM group.

ICU transfer:

4 studies stated (ICU transfer), & all can be applied. A significant heterogeneity was found. Consequently, a random-influence model was applied for analysis ($I^2 = 80\%$, p-value equal 0.002). The combined odds ratio and 95% CIs were 0.90 (0.76 to 1.07). The combined finding determines a statistically insignificant difference among groups according to ICU transfer (Z equal to 1.24, P equal to 0.22).

Study or Subgroup	MLM group Events	MLM group Total	No-MLM group Events	No-MLM group Total	Weight	Odds Ratio M-H, Random, 95% CI
Bailey, 2013	444	9911	426	10120	28.6%	1.07 [0.93, 1.22]
Bassin, 2023	104	14738	132	14352	19.5%	0.77 [0.59, 0.99]
Escobar, 2020	2349	13274	4974	23797	33.8%	0.81 [0.77, 0.86]
Levin, 2024	115	1488	100	1262	18.1%	0.97 [0.74, 1.29]
Total (95% CI)		**39411**		**49531**	**100.0%**	**0.90 [0.76, 1.07]**
Total events	3012		5632			

Heterogeneity: Tau² = 0.02; Chi² = 14.74, df = 3 (P = 0.002); I² = 80%
Test for overall effect: Z = 1.24 (P = 0.22)

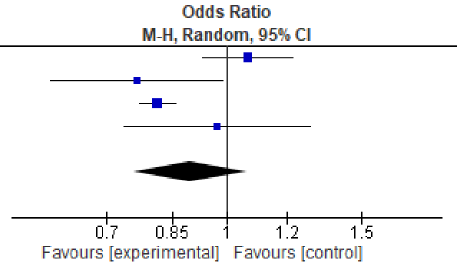

Figure (2): Forest plot of ICU transfer reveals a statistically insignificant variance among MLM group and the No-MLM group.

Length of hospital stays:

Four studies (length of hospital stays) stated that all may be applied. A significant heterogeneity was found. Thus, a random-influence model was applied for analysis ($I^2 = 95\%$, p below 0.00001). The combined mean distinction and ninety-five percent CIs were -0.35 (-0.68 to -0.01). The combined finding reveals a statistically significant difference between groups in terms of length of hospital stays (Z equal to 2.01, P equal to 0.04).

Study or Subgroup	MLM group Mean	MLM group SD	MLM group Total	No-MLM group Mean	No-MLM group SD	No-MLM group Total	Weight	Mean Difference IV, Random, 95% CI
Bailey, 2013	3.9	3.55	9911	3.82	3.58	10120	27.6%	0.08 [-0.02, 0.18]
Bassin, 2023	4.307	4.0183	14540	4.56	4.4483	14099	27.6%	-0.25 [-0.35, -0.15]
Escobar, 2020	6.5	8	13274	7.2	10.1	23797	26.0%	-0.70 [-0.89, -0.51]
Levin, 2024	6.6203	5.5207	1488	7.234	6.3158	1252	18.8%	-0.61 [-1.06, -0.17]
Total (95% CI)			39213			49268	100.0%	-0.35 [-0.68, -0.01]

Heterogeneity: Tau² = 0.10; Chi² = 61.97, df = 3 (P < 0.00001); I² = 95%
Test for overall effect: Z = 2.01 (P = 0.04)

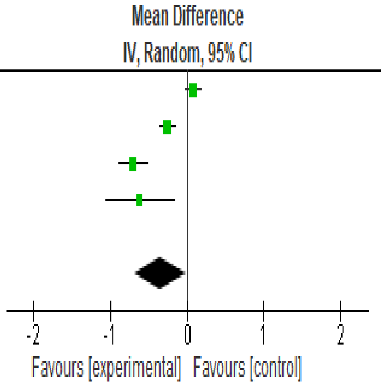

Figure (3): Forest plot of Length of hospital stays demonstrates a statistically significant variance among the MLM group and the No-MLM group.

5. DISCUSSION

Adults experiencing deterioration in general medical-surgical wards face significant mortality and morbidity. Initiatives for the early recognition of clinical worsening in inpatients outside the ICU have utilized manually computed scores, like the National Early Warning Score, which necessitates manual chart abstraction of essential signs and point allocation depending on these metrics; should a case's score surpass a designated threshold, a rapid-response team is summoned. **(13, 14).**

Certain research has delineated automated essential-sign triggers and automated iterations of the National Early Warning Score or similar scores. Several researchers have created intricate predictive models, appropriate for real-time utilization with electronic health records (EHRs), aimed at the early identification of patient condition worsening, involving one model that underwent testing in a randomized trial. These electronic health records-based models involve laboratory tests and data regarding comorbidities; they may entail intricate calculations. **(15, 16).**

AI-EWS for clinical deterioration is showing significant promise in improving patient outcomes by providing more accurate and timely alerts compared to traditional methods, leading to earlier interventions and a reduction in adverse events like cardiac arrest, ICU, and mortality **(17).**

Studies indicate these AI systems can identify cases at high risk of decline, enabling care teams to intervene sooner, which is correlated with lower mortality rates and reduced hospital stays. While further real-world research and clinician adherence are necessary, AI-EWS represents a powerful and cost-effective tool for enhancing case safety and quality of care **(18, 19).**

This meta-analysis included (5) studies **(8-12).**
Our results revealed that the mean age in the examined groups was 62.3, ranging from 25 to 76 years, and gender has been stated in 2 studies, with 29604 males and 27498 females.

Our results, supported by **Bailey TC et al. (12),** stated that the median age of the studied cases was fifty-seven years, with a range of forty-four to sixty-nine years in the control group and fifty-seven years, with a range of forty-four to seventy years, in the intervention group. Otherwise, there were 53% females in the control group and 54% were females in the intervention group.

Also, **Escobar GJ et al**. **(11)** reported that the mean age of the examined cases in the intervention group was 68.4±15.6 years and there were 55.1% men. The mean age of the examined cases in the Comparison Cohort was 67.2±16.2 years, and there were 54.3% males.

Regarding mortality, our results showed that significant heterogeneity was observed. Thus, a random-influence model was applied for analysis ($I^2 = 85\%$, p below 0.0001). The combined odds ratio and ninety-five percent CIs were 0.76 (0.63 to 0.91). The combined finding reveals a statistically significant difference between groups according to mortality (Z equal to 2.89, P equal to 0.004).

Our results agreed with **Escobar GJ et al. (11), who** reported that mortality within thirty days following an alert was reduced in the intervention group than in the comparison group (adjusted relative risk, 0.84; ninety-five percent CI, 0.78 to 0.90; P below 0.001).

In addition, **Bassin L et al. (9)** reported that the intervention group illustrated decreased probability for every finding barring death. At the event level, there were significantly fewer MAEs [5.7 percent (852/14,738)] against [7.1 percent (1,024/14,352), P below 0.001] in the intervention group.

Moreover, **Levin MA et al. (8)** reported that combined in-hospital and thirty-day mortality decreased in the intervention group.

Furthermore, **Winslow CJ et al**. **(10)** stated that there was a significant variance among the studied groups according to mortality.

On the other hand, **Bailey TC et al. (12)** stated that there were no variances in the proportion of cases who died in the intervention group than the control group.

In the present research, significant heterogeneity has been observed regarding ICU transfer. Thus, a random-influence model was utilized for analysis ($I^2 = 80\%$, $p = 0.002$). The combined odds ratio and 95% CIs were 0.90 (0.76 to 1.07). The combined finding demonstrates a statistically insignificant difference among groups according to ICU transfer (Z=1.24, P equal to 0.22).

A randomized, controlled crossover study conducted by **Bailey TC et al. (12)** reported that no variances have been observed in the proportion of cases who have been transferred to the ICU in the intervention group than the control group.

In the USA, **Escobar GJ et al. (11)** reported that cases in the intervention group, compared to those in the comparison group, had a decrease in unadjusted frequency of ICU admission (17.7% against 20.9%).

In Australia, **Bassin L et al. (9)** reported that admissions via the emergency department (ED) were significantly fewer in the intervention against the control group [40.4 percent (5,871/14,540) against 41.6 percent (5,871/14,099), P equal to 0.03].

We found that significant heterogeneity was detected regarding the length of hospital stays. Therefore, a random-influence model was applied for analysis ($I^2 = 95\%$, p below 0.00001). The combined mean distinction and 95% CIs were -0.35 (-0.68 to -0.01). The combined finding demonstrates a statistically significant variance among groups according to length of hospital stays (Z equal to 2.01, P equal to 0.04).

This came in concordance with **Escobar GJ et al. (11), who** reported that cases in the intervention group than those in the comparison group, had shorter lengths of stay between survivors (6.5 days versus 7.2 days).

Meanwhile, **Bassin L et al**. **(9)** reported that the intervention group had a significantly shorter hospitalization period (3.74 days, IQR 1.84–7.26) than the control group (3.86 days, IQR 1.86–7.86, P equal to 0.002).

However, **Bailey TC et al. (12)** reported that no variance has been found in LOS in the intervention group than the control group.

6. CONCLUSION

The results indicate that AI-driven early warning systems enhance patient outcomes in practical clinical environments. Notwithstanding the apparent advantages, the efficacy and practical usability of these models necessitate additional investigation.

REFERENCES

[1] Yuan S, Yang Z, Li J, Wu C, Liu S. AI-Powered early warning systems for clinical deterioration significantly improve patient outcomes: a meta-analysis. BMC Medical Informatics and Decision Making. 2025 Jun 2;25(1):203.
[2] Veldhuis LI, Woittiez NJ, Nanayakkara PW, Ludikhuize J. Artificial intelligence for the prediction of in-hospital clinical deterioration: a systematic review. Critical care explorations. 2022 Sep 1;4(9): e0744.
[3] Mann KD, Good NM, Fatehi F, Khanna S, Campbell V, Conway R, Sullivan C, Staib A, Joyce C, Cook D. Predicting patient deterioration: a review of tools in the digital hospital setting. Journal of medical Internet research. 2021 Sep 30;23(9):e28209.
[4] Jahandideh S, Ozavci G, Sahle BW, Kouzani AZ, Magrabi F, Bucknall T. Evaluation of machine learning-based models for prediction of clinical deterioration: A systematic literature review. International Journal of Medical Informatics. 2023 Jul 1; 175:105084.
[5] Muralitharan S, Nelson W, Di S, McGillion M, Devereaux PJ, Barr NG, Petch J. Machine learning–based early warning systems for clinical deterioration: systematic scoping review. Journal of medical Internet research. 2021 Feb 4;23(2):e25187.
[6] Shamout FE, Shen Y, Wu N, Kaku A, Park J, Makino T, Jastrzębski S, Witowski J, Wang D, Zhang B, Dogra S. An artificial intelligence system for predicting the deterioration of COVID-19 patients in the emergency department. NPJ digital medicine. 2021 May 12;4(1):80.
[7] Peng J, Zou K, Zhou M, Teng Y, Zhu X, Zhang F, Xu J. An explainable artificial intelligence framework for the deterioration risk prediction of hepatitis patients. Journal of medical systems. 2021 May;45(5):61.
[8] Levin MA, Kia A, Timsina P, et al. Real-Time machine learning alerts to prevent escalation of care: a nonrandomized clustered pragmatic clinical trial. Crit Care Med. 2024;52(7):1007–20.
[9] Bassin L, Raubenheimer J, Bell D. The implementation of a real time early warning system using machine learning in an Australian hospital to improve patient outcomes. Resuscitation. 2023;188:109821.

[10] Winslow CJ, Edelson DP, Churpek MM, et al. The impact of a machine learning early warning score on hospital mortality: A multicenter clinical intervention trial. Crit Care Med. 2022;50(9):1339–47.

[11] Escobar GJ, Liu VX, Schuler A, et al. Automated identification of adults at risk for In-Hospital clinical deterioration. N Engl J Med. 2020;383(20):1951–60.

[12] Bailey TC, Chen Y, Mao Y, et al. A trial of a real-time alert for clinical deterioration in patients hospitalized on general medical wards. J Hosp Med. 2013;8(5):236–42

[13] Chung J, Jung H. Indicators of clinical deterioration in adult general ward patients from nurses' perspectives: a mixed-methods systematic review. BMC nursing. 2024 Nov 26;23(1):861.

[14] Henry KE, Giannini HM. Early warning systems for critical illness outside the intensive care unit. Critical care clinics. 2024 Jul 1;40(3):561-81.

[15] Ruiz VM, Goldsmith MP, Shi L, Simpao AF, Gálvez JA, Naim MY, Nadkarni V, Gaynor JW, Tsui FR. Early prediction of clinical deterioration using data-driven machine-learning modeling of electronic health records. The Journal of thoracic and cardiovascular surgery. 2022 Jul 1;164(1):211-22.

[16] Fu LH, Knaplund C, Cato K, Perotte A, Kang MJ, Dykes PC, Albers D, Collins Rossetti S. Utilizing timestamps of longitudinal electronic health record data to classify clinical deterioration events. Journal of the American Medical Informatics Association. 2021 Sep 1;28(9):1955-63.

[17] Villanueva-Miranda I, Xiao G, Xie Y. Artificial Intelligence in Early Warning Systems for Infectious Disease Surveillance: A Systematic Review. Frontiers in Public Health. 2025 Jun 23;13:1609615.

[18] Shaik TB. Revolutionizing healthcare with federated reinforcement learning: from machine learning to machine unlearning.

[19] Shah PK, Ginestra JC, Ungar LH, Junker P, Rohrbach JI, Fishman NO, Weissman GE. A simulated prospective evaluation of a deep learning model for real-time prediction of clinical deterioration among ward patients. Critical care medicine. 2021 Aug 1;49(8):1312-21.