Prehospital and Disaster Medicine

www.cambridge.org/pdm

Quantitative Triage and Cumulative Risk in Neonatal Intensive Care Unit Evacuation: A Simulation Study Using Real-World Clinical Data

Kazunori Imai MD, PhD, MBA[1,2], Chunxi Xu MIS[3], Shoya Oikawa MIS[3], Yuko Araki PhD[3], Shinji Saitoh MD, PhD[2] and Osuke Iwata MD, PhD[2]

[1]Department of Acute Care and Disaster Medicine, Nagoya City University Graduate School of Medical Sciences and Medical School, Nagoya, Japan; [2]Department of Pediatrics and Neonatology, Nagoya City University Graduate School of Medical Sciences and Medical School, Nagoya, Japan and [3]Graduate School of Information Sciences, Tohoku University, Sendai, Japan

Original Research

Cite this article: Imai K, Xu C, Oikawa S, Araki Y, Saitoh S, and Iwata O. (2026) Quantitative Triage and Cumulative Risk in Neonatal Intensive Care Unit Evacuation: A Simulation Study Using Real-World Clinical Data. *Prehospital and Disaster Medicine*. **41**(e27), 1–8. doi: 10.1017/S1049023X26109157

Received: 23 January 2026 UTC
Revised: 4 July 2026 UTC
Accepted: 2 September 2026 UTC

Keywords:
Disaster Planning; Emergency Medical Services; Infant, Newborn; Intensive Care Units, Neonatal; Triage

Abbreviations:
NEXT, Neonatal Extrication Triage; NICU, Neonatal Intensive Care Unit; NTISS, Neonatal Therapeutic Intervention Scoring System; START-Neo-R, Simple Triage and Rapid Treatment for Neonates, Revised; TES, Total Evacuation Score; TRAIN, Triage by Resource Allocation for IN-patients

Corresponding author:
Kazunori Imai
Department of Acute Care and Disaster Medicine
Nagoya City University, Graduate School of Medical Sciences
1-Kawasumi, Mizuho-cho, Mizuho-ku
Nagoya, Aichi, JAPAN, ZIP 467-8602
E-mail: k.imai@med.nagoya-cu.ac.jp

Abstract

Introduction: Neonatal intensive care unit (NICU) evacuation poses distinct operational and ethical challenges because neonates often depend on continuous life-sustaining technologies and specialized care. Although several neonatal or pediatric disaster triage systems have been proposed, actual NICU evacuation cannot be ethically or practically tested under controlled disaster conditions. A reproducible framework comparing evacuation-order strategies is therefore needed.

Study Objective: This study aimed to develop and apply a reproducible simulation framework for NICU evacuation-order evaluation and, within that framework, to compare structured triage with evacuation without triage and quantitative with categorical neonatal triage approaches.

Methods: A retrospective simulation study was conducted using routinely recorded clinical triage assessments from a tertiary NICU, comprising 1,079 patient evaluations in 162 neonates across 49 triage days. During clinical practice, clinical severity and triage status had been recorded using the Neonatal Therapeutic Intervention Scoring System (NTISS), Neonatal Extrication Triage (NEXT), and Simple Triage and Rapid Treatment for Neonates, Revised (START-Neo-R); NEXT was a quantitative six-domain resource-dependency score ranging from zero to 12, whereas START-Neo-R was a five-level categorical neonatal triage system. Three evacuation strategies were evaluated: descending-order evacuation using NEXT, descending-order evacuation using START-Neo-R, and evacuation without triage. Evacuation-order performance was quantified using the Total Evacuation Score (TES), defined as the time integral of cumulative NTISS scores among patients remaining in the NICU across sequential evacuation steps. The TES distributions were compared descriptively across strategies rather than inferentially.

Results: Within the 49 observed triage days, both structured triage strategies generated TES distributions shifted lower than those from evacuation without triage. In the direct comparison between NEXT and START-Neo-R, the worst-case TES under NEXT was lower than more than one-half of the START-Neo-R TES distribution on 34 of the 49 observed triage days (69.4%).

Conclusion: In this simulation study using real-world NICU data, TES provided a reproducible framework for comparing neonatal evacuation-order strategies under explicit assumptions. Within this TES-based framework, structured triage generated TES distributions shifted lower than those from evacuation without triage, and NEXT showed a more favorable distribution-based pattern than START-Neo-R on most observed days. These findings suggest that TES-based evaluation may help identify triage strategies that support more transparent and testable NICU evacuation planning.

© The Author(s), 2026. Published by Cambridge University Press on behalf of World Association for Disaster and Emergency Medicine. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Introduction

Large-scale disasters, including climate-related hazards, increasingly threaten health systems world-wide.[1] These events expose gaps in systems designed to protect vulnerable populations. Children, particularly newborns, face unique and disproportionate risks during disasters. As the American Academy of Pediatrics (Itasca, Illinois USA) has emphasized, children are not "small adults;" their anatomical, physiological, and psychological characteristics require pediatric-specific disaster planning and response strategies.[2] Recent disasters, including Hurricane Sandy in the United States in 2012 and the 2016 Kumamoto Earthquake in Japan, have demonstrated that neonatal and pediatric evacuations pose substantial challenges under resource-constrained conditions.[3,4]

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 08:07:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1049023X26109157

Table 1. Neonatal Extrication Triage

Category	Score		
	0	1	2
Body Weight and Gestational Age	≥2,500g and ≥37 weeks	≥1,000g to <2,500g OR ≥28 weeks to <37 weeks	<1,000g or <28 weeks
Respiratory Support	None or oxygen supplementation	On HFNC or NIPPV	Using invasive ventilation
Indwelling Cannula (excluding endotracheal tube)	None	≤3	>3
Monitoring	None	Pulse oximeter	Pulse oximeter and ECG
Incubator	Open bed	Closed incubator with ambient temperature <30°C or Open incubator without heating	Closed incubator with ambient temperature ≥30°C or Open incubator with heating
Expected Risk when Monitoring/Treatments are Disrupted during Transportation	None or minor (not life-threatening)	Moderate (possibly leading to deteriorating clinical conditions)	Major (possibly leading to life-threatening events)

Note: A score of 0–2 was assigned for each of the six domains to give a composite score of 0–12.
Abbreviations: ECG, electrocardiogram; HFNC, high-flow nasal cannula; NIPPV, noninvasive positive-pressure ventilation.

These challenges are most acute in neonatal intensive care units (NICUs), where critically ill neonates depend on continuous access to life-sustaining technologies.[5] Hospital-wide evacuation requires coordination across departments with competing priorities. Without a shared, objective triage framework, neonatal patients may be under-valued or deprioritized when rapid decisions are required under resource constraints.[5,6] Effective disaster planning therefore requires triage systems that can guide evacuation priorities based on clinical urgency, technology dependence, or specialized resource requirements.

The appropriate evacuation order may vary by disaster context.[3–5] When a hospital is relatively safe but losing resources, evacuating the sickest patients first may allow more severely ill neonates to continue intensive care at an undamaged center. Conversely, in time-critical events such as fire or structural collapse, least sick first evacuation may maximize rapid clearance from the damaged hospital. Disaster coordinators must therefore choose evacuation priorities according to residual hospital function, time pressure, transport capacity, and receiving capacity.

Several neonatal or pediatric in-patient triage systems have been proposed, including the five-level Simple Triage and Rapid Treatment for Neonates, Revised (START-Neo-R; Supplementary Material 1) and Triage by Resource Allocation for IN-patients (TRAIN).[7,8] However, experience from the 2016 Kumamoto Earthquake suggested that applying START-Neo-R in a NICU could concentrate many patients within a single category, leaving limited granularity for deciding evacuation order.[4,9] In response to this problem, Neonatal Extrication Triage (NEXT) was developed as a quantitative composite score for neonatal medical resource dependency (Table 1).[9] In its development study, NEXT showed criterion-related validity against the Neonatal Therapeutic Intervention Scoring System (NTISS)[10] and high inter-rater reliability among NICU physicians and nurses, while providing finer patient dispersion than START-Neo-R.[9]

Even with these triage tools, quantitative evaluation of evacuation triage has been limited by the absence of a standardized method for comparing evacuation-order strategies. Most prior work has focused on conceptual development, expert consensus, feasibility, or reliability rather than reproducible quantitative validation of evacuation order. This gap is understandable: actual NICU evacuation cannot be ethically studied under controlled conditions, and real disaster events are rare, heterogeneous, and strongly influenced by local resources, transfer routes, staffing, equipment, and receiving capacity.

To address this methodological gap, the Total Evacuation Score (TES) was developed as a distribution-based metric linking evacuation order to the cumulative burden of illness among patients remaining in the NICU. In this study, TES was constructed using NTISS, a validated therapy-based neonatal severity index;[10] NTISS was selected as the reference severity measure because no universally accepted neonatal evacuation-priority standard was available and because it had been used to assess the criterion-related validity of NEXT in the preceding development study.[9] Its therapy-based structure captures treatment intensity and technology dependence, which are directly relevant to evacuation prioritization.[10] The TES framework itself is not limited to NTISS and could be adapted to other locally or nationally accepted severity measures.

Within this framework, the present simulation study used real-world NICU patient data to evaluate two structured triage strategies: NEXT and START-Neo-R. Because descending and ascending evacuation orders are reciprocal when each evacuation step is assigned an equal interval, this study focused on descending evacuation as an analytically tractable baseline. The aims were to develop and apply a reproducible simulation framework for NICU evacuation-order evaluation, to compare structured triage with evacuation without triage, and to compare NEXT with START-Neo-R as quantitative and categorical triage approaches, respectively. As a simulation-based approach, the study was intended to support transparent and reproducible comparisons of evacuation-order strategies under explicit assumptions.

Methods

Study Design and Setting

This retrospective simulation study used a dataset assembled from routinely recorded clinical triage assessments that had been used to

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 08:07:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1049023X26109157

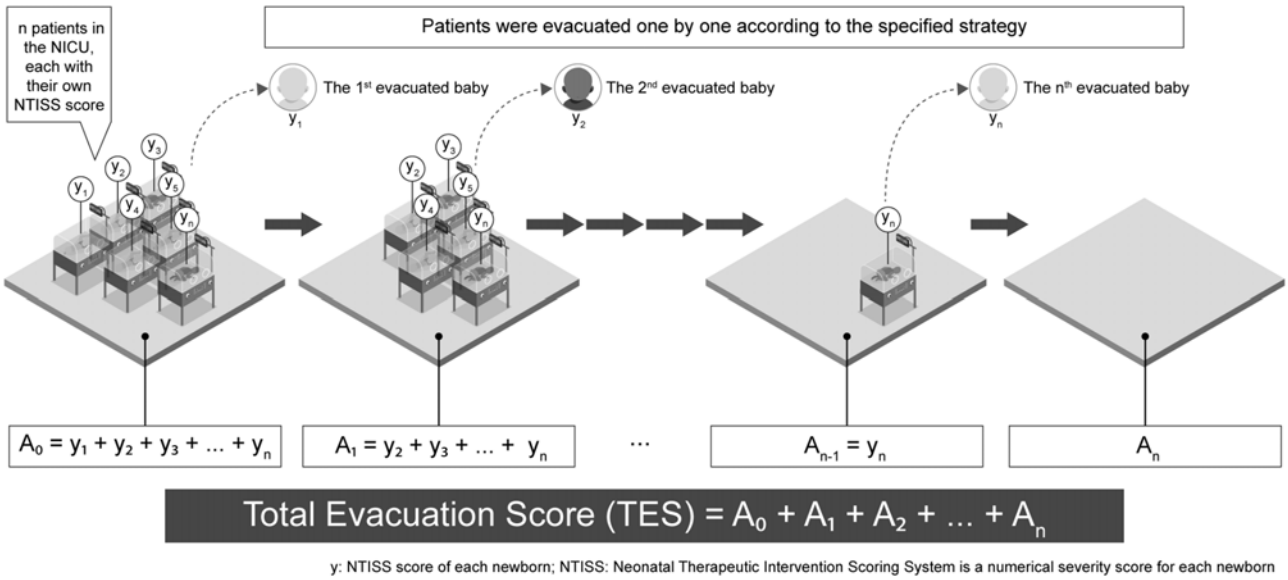


Figure 1. Definition of the Total Evacuation Score (TES).
Note: The TES measures the cumulative severity of illness of patients remaining in the neonatal intensive care unit (NICU) during a simulated evacuation sequence. It is calculated by summing the total Neonatal Therapeutic Intervention Scoring System (NTISS) scores of all patients remaining in a particular unit at each sequential evacuation step. Lower TES values indicate earlier positioning of higher-acuity patients in the evacuation order under the equal-step assumption. Abbreviations: NICU, neonatal intensive care unit; NTISS, Neonatal Therapeutic Intervention Scoring System; TES, Total Evacuation Score.

develop and evaluate the NEXT system.[9] The source data were obtained from clinical assessments conducted from August 1, 2019 through February 28, 2020 at a single tertiary NICU in a university hospital in Japan. During routine clinical practice, NEXT, START-Neo-R, and NTISS had been recorded before the present simulation analysis was conducted.

Patient Cohort and Data Collection

Detailed characteristics of this patient cohort have been previously published.[9] At each assessment point, a physician evaluated each neonate using NEXT and START-Neo-R and calculated NTISS.

This study was approved by the Institutional Review Board (IRB) of Nagoya City University (Nagoya, Aichi Prefecture, Japan; Approval number: 60-20-0110). The requirement for informed consent was waived by the IRB owing to the retrospective nature of the analysis and anonymization of data in accordance with institutional ethical standards and the 1964 Declaration of Helsinki and its later amendments.

Triage Systems and Severity Assessment

The NEXT triage system is a quantitative tool designed to assess dependency on medical resources.[9] It evaluates six domains, including immaturity, respiratory support, and indwelling cannula, and assigns a score of zero-to-two to each. The sum of these scores yields a composite score ranging from zero to 12 (Table 1).

The START-Neo-R triage system is a five-level color-coded categorical system (Green, Yellow, Red-I, Red-II, and Gray) that classifies neonates based on clinical severity and care dependency (Supplementary Material 1).[7]

The NTISS tool is a therapy-based neonatal severity assessment tool comprising 64 items, with higher scores indicating greater illness severity.[10] It had been recorded in the source dataset and had been used as the reference severity measure when assessing the criterion-related validity of NEXT because it is a validated neonatal severity index and captures the intensity of therapeutic interventions relevant to evacuation prioritization.[9,10] Unlike adult critical care, neonatal intensive care does not have a single universally adopted severity score comparable to commonly used adult organ-failure scores. The NTISS tool was not assumed to represent a direct measure of evacuation outcome, transport complexity, or survival benefit.

Evacuation Simulation and Efficacy Metric

Evacuation simulations were conducted using a single pre-defined evacuation direction. Both descending (sickest first) and ascending (least sick first) evacuation strategies were considered conceptually because the appropriate direction may differ by disaster context. Under the simplifying assumption that each evacuation step has an equal interval, descending and ascending orders are mathematically reciprocal. Accordingly, descending evacuation was analyzed as a controlled baseline to isolate the effect of prioritization order.

To quantify evacuation-order performance, TES was defined as the time integral of cumulative NTISS severity among patients remaining in the NICU throughout the simulated evacuation process. Under the equal-step assumption, this time integral was represented as the discrete sum of the remaining NTISS burden after each evacuation step. Specifically, TES was calculated as:

$$TES = \Sigma\, Ai, from\, i = 0\, to\, n,$$

where Ai is the total NTISS score of patients remaining after the i-th evacuation (Figure 1).

Lower TES values indicated that patients with higher NTISS scores were positioned earlier in the evacuation sequence.

The equal-step assumption was used to isolate evacuation order from site-specific operational variables. In actual evacuation, transport time and staffing requirements may differ according to respiratory support, hemodynamic instability, infusion pumps, incubator requirements, transport route, transfer distance, and receiving capacity. Because the source dataset did not contain

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 08:07:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1049023X26109157

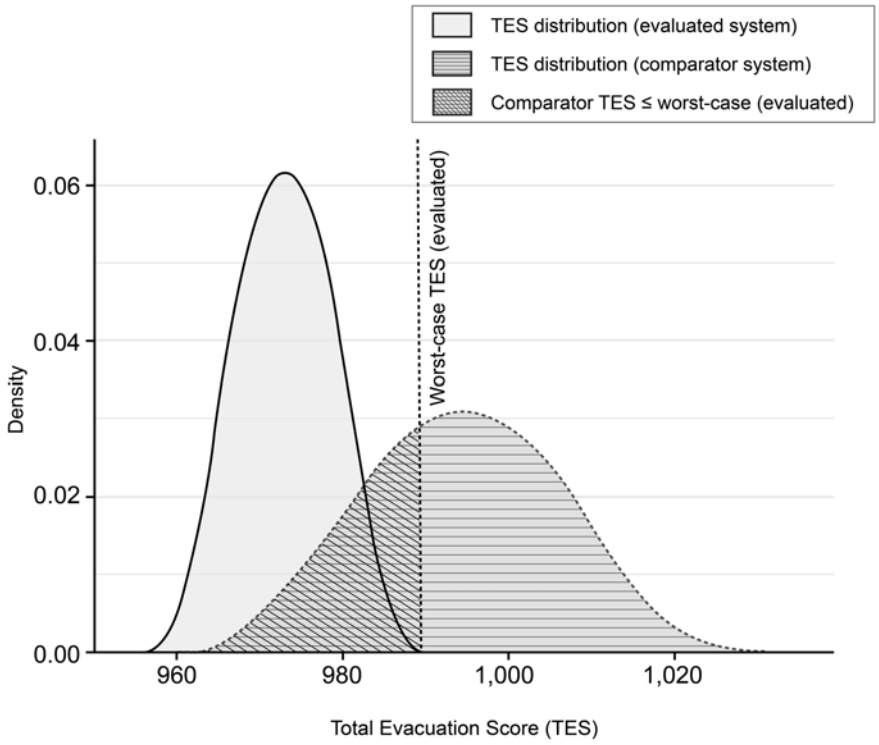


Figure 2. Comparison of Total Evacuation Score (TES) Distributions.
Note: The TES distributions for an evaluated strategy and a comparator are shown. The vertical line indicates the evaluated strategy's worst-case TES. The highlighted portion of the comparator distribution at or below this line represents the pre-defined shaded-area probability: the proportion of comparator TES values that are equal to or more favorable than the evaluated strategy's worst-case TES. The evaluated strategy is considered favorable under the pre-defined distribution-based criterion if the pre-defined shaded-area probability is <0.5. Abbreviation: TES, Total Evacuation Score.

measured data on transport time, staffing burden, or other operational factors, assigning weights to these factors would have required post hoc assumptions; therefore, such weights were not included in the primary model.

Simulations were performed for three strategies: (1) evacuation based on descending NEXT scores; (2) evacuation based on descending START-Neo-R categories; and (3) the no-triage strategy, defined as evacuation without triage, with patient order randomly assigned.

Statistical Analysis

For each observed triage day, TES distributions were analyzed separately. Because TES values vary according to the number and acuity of patients present on each day, direct pooling of TES values across days was not appropriate. Instead, each strategy was compared within each observed triage day.

For NEXT and START-Neo-R, all possible permutations of patient order within tied scores or categories were exhaustively enumerated to generate the complete TES distribution under each rule. For the no-triage strategy, the total number of possible sequences was computationally prohibitive (up to 10^{28}), so random sampling with replacement was used to approximate the no-triage permutation distribution. For this approximation, the relevant sampling uncertainty was the accuracy of approximating the full permutation distribution, rather than uncertainty around a population-level effect estimate. The sampling procedure was validated by confirming that sampled TES distributions reproduced the mean and variance of complete permutation distributions in settings where complete enumeration was feasible. Details are provided in Supplementary Material 2.

The TES distributions of each system were visualized side-by-side as histograms (Figure 2). A vertical dashed line was drawn at the worst-case TES of the evaluated triage system. The pre-defined shaded-area probability was defined as the proportion of comparator TES values on or below this line, representing comparator TES values that were equal to or more favorable than the evaluated system's worst-case TES. The threshold of 0.5 was specified a priori as a conservative distribution-based benchmark: the evaluated strategy was considered favorable only when the pre-defined shaded-area probability was <0.5, meaning that worst-case TES was lower than more than one-half of the comparator distribution. This threshold was not intended as a clinically validated cutoff.

Because the objective was to describe TES distributions generated by pre-defined evacuation-order rules within the observed dataset, the primary analysis was descriptive and distribution-based rather than inferential. The study did not estimate an intervention effect or a population-level association; therefore, conventional confidence intervals and effect sizes were not pre-specified as primary summaries. Results were instead reported using distribution-based descriptive summaries, including the number of observed triage days on which the pre-defined shaded-area probability criterion was met, the number of days with the maximum unfavorable value of 1.0, day-level values, and the relative dispersion of TES distributions.

Results

Patient Characteristics

The study cohort comprised 1,079 patient evaluations of 162 neonates across 49 simulation days. Baseline scores included a median NTISS score of six (interquartile range [IQR], 3-13) and a median NEXT score of six (IQR, 3-8). Patients were classified by START-Neo-R as follows: Green, 30.2%; Yellow, 19.0%; Red-I, 23.0%; Red-II, 25.3%; and Gray, 2.5%.

Simulation Parameters

The number of simulated evacuation patterns varied according to patient census and triage score distribution across the 49 days. For the no-triage strategy, the random sampling approach produced

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 08:07:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1049023X26109157

between 8.2×10^5 and 3.3×10^8 simulations per day. For the structured triage strategies, the number of unique evacuation sequences ranged from 2,304 to 5.8×10^7 for NEXT and from 5.2×10^7 to 3.3×10^{13} for START-Neo-R. The random sampling method for the no-triage strategy was validated by confirming consistency between sampled and full-set TES distributions where enumeration was feasible (Supplementary Material 2).

Comparison of Triage Efficacy

Within the observed dataset, both structured triage systems generated more favorable TES distributions than evacuation without triage. When each structured triage system was evaluated against the no-triage strategy as the comparator, the TES distributions for both NEXT and START-Neo-R were lower than those for the no-triage strategy on all 49 simulation days, with complete separation of the distributions. Under the assumptions of the simulation model, even the least favorable triage-guided sequence within tied scores or categories produced a lower TES than the sampled no-triage evacuation sequences (Figure 3A and Supplementary Material 3).

When NEXT was evaluated against START-Neo-R as the comparator, the pre-defined shaded-area probability was below the 0.5 threshold on 34 of the 49 observed triage days (69.4%). This condition was not satisfied on the remaining 15 observed triage days. The pre-defined shaded-area probability was 1.0 on five of the 49 observed triage days, indicating days on which all TES values under NEXT were higher than those under START-Neo-R in this descriptive comparison. Day-level pre-defined shaded-area probabilities for all 49 days are provided in Supplementary Material 4.

The TES distributions were classified into five representative patterns based on the degree of overlap between NEXT and START-Neo-R (Figure 3B-F). Full TES distribution plots for all 49 days are presented in Supplementary Material 5.

Discussion

Interpretation

Previous neonatal and pediatric disaster triage studies have largely focused on tool development, feasibility, reliability, or expert consensus rather than quantitative evaluation of evacuation-order performance.[7–9] The principal contribution of this study is the development and application of a reproducible simulation framework for evaluating NICU evacuation order using observed patient-level clinical data. Although this simulation cannot show whether a triage strategy improves patient outcomes during an actual disaster, TES makes it possible to examine whether simulated evacuation orders align with a pre-specified measure of patient burden. A more favorable TES distribution means that patients with greater measured burden are less likely to remain in the NICU until later evacuation steps under the modeled order. In this study, NTISS was used as that reference measure, but the TES concept is not inherently NTISS-specific and could be recalculated using other locally accepted severity or transport-burden measures.

Within this framework, structured triage generated more favorable TES distributions than evacuation without triage. In the comparison between NEXT and START-Neo-R, the pre-defined shaded-area probability was below the 0.5 threshold on 34 of the 49 observed triage days. The TES histograms also suggested more concentrated distributions for NEXT than for START-Neo-R, consistent with the greater granularity of its quantitative structure. These findings suggest that structured triage may support more favorable evacuation-order planning than evacuation without triage and that, when structured triage is used, NEXT may provide more reproducible prioritization than START-Neo-R.

Comparison of Triage Systems

The observed pattern favoring NEXT over START-Neo-R can be understood in relation to score granularity and alignment with NTISS burden within the TES framework; NEXT more often generated evacuation orders aligned with higher NTISS burden, whereas START-Neo-R groups diverse patients into a smaller number of categories, creating ties and ambiguity in prioritization. This categorical structure may widen TES distributions because the order within each category remains less constrained. Conversely, NEXT provides greater granularity, reduces ties, and generates a more constrained evacuation hierarchy.

The degree of alignment with NTISS burden varied across observed triage days. A one-point change in NEXT or one-category change in START-Neo-R does not necessarily correspond to a fixed increment in NTISS, so the relative performance of the two systems may depend on the patient mix present on a given day. This variability supports evaluating triage systems across multiple observed census patterns rather than relying on a single representative scenario.

Other in-patient or pediatric disaster triage systems, such as TRAIN, address related questions of resource allocation and prioritization.[8] Additional systems were not applied retrospectively because the variables required to assign them reliably were not recorded in a standardized manner, and post hoc classification without those variables could introduce misclassification bias. Future studies should prospectively collect the variables needed to compare multiple neonatal and pediatric evacuation triage systems within the same TES-based framework.

Assumptions and Future Extensions of the TES Framework

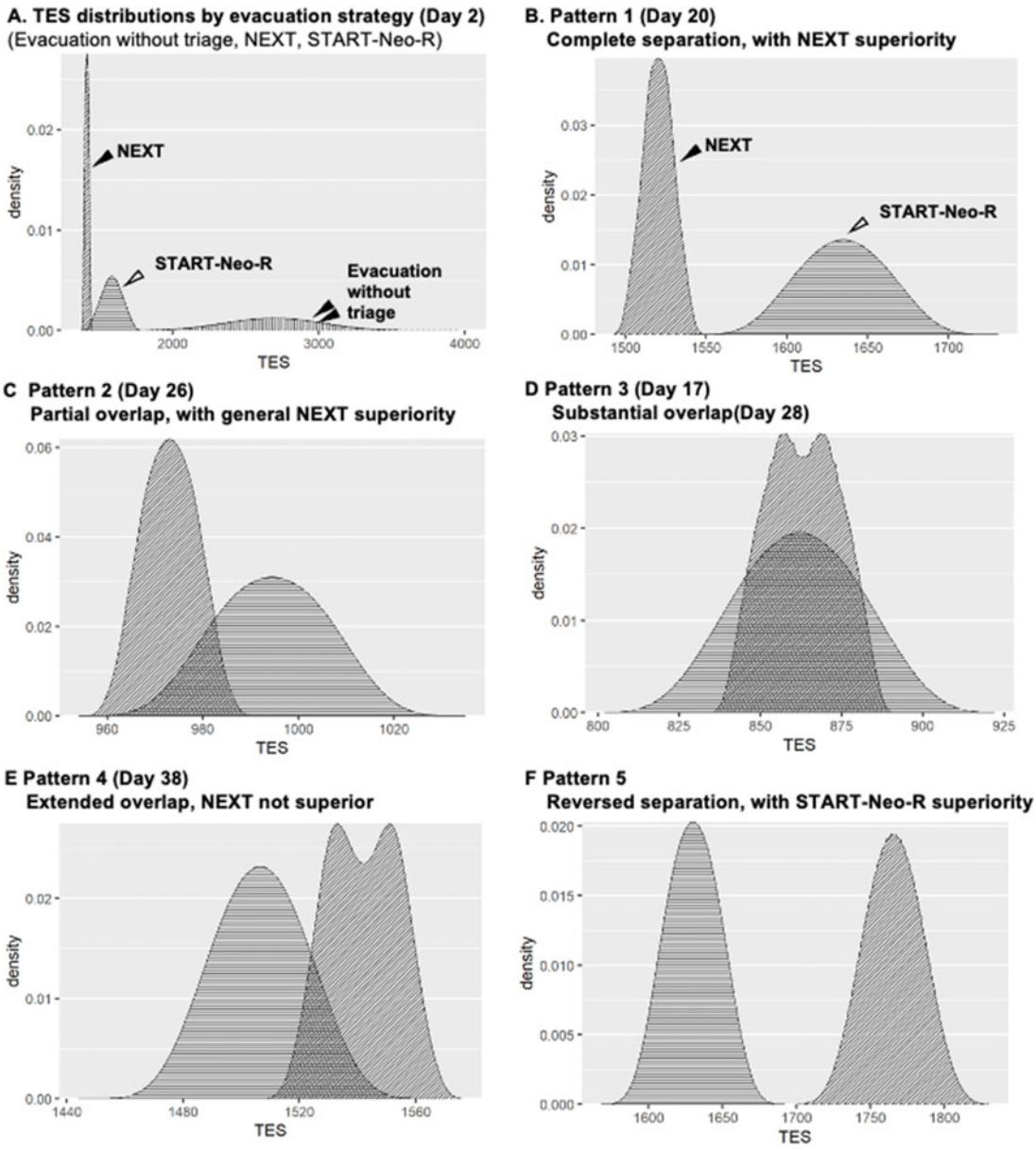
The TES framework was intentionally implemented in its simplest order-based form to isolate the effect of evacuation order from site-specific operational factors. This study simulated descending evacuation, a strategy that may be relevant when hospital function is deteriorating but downstream transport and receiving capacity remain partially preserved. Time-critical disasters, such as fire or structural collapse, may instead require least-sick first evacuation to maximize speed and clearance. Under the equal-step assumption, descending and ascending evacuation orders are reversible, allowing the same framework to evaluate alternative evacuation directions.

This simple formulation also provides a foundation for future weighted models. Mechanically ventilated, hemodynamically unstable, or technology-dependent neonates may require more time, staff, equipment, and coordination than stable patients, but the degree of additional burden is likely to vary across institutions. Rather than assigning universal weights without empirical data, future studies could extend TES by deriving patient-or site-specific weights from evacuation drills, transport-time measurements, staffing experiments, or real evacuation records.

Clinical and Disaster-Preparedness Implications

The TES framework may help NICUs translate patient-level clinical information into a transparent evacuation-order plan and

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 08:07:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1049023X26109157



G Summary of TES distribution patterns and triage days

Distribution Type	Number of Days	Triage day
Pattern 1	3	1, 15, 40
Pattern 2	19	2, 4, 5, 6, 12, 13, 16, 20, 24, 25, 30, 31, 32, 33, 39, 46, 47, 48, 49
Pattern 3	16	3, 7, 8, 9, 11, 18, 19, 21, 22, 26, 34, 41, 42, 43, 44, 45
Pattern 4	6	10, 14, 27, 28, 29, 36
Pattern 5	5	17, 23, 35, 37, 38

Figure 3. Distribution Patterns of Total Evacuation Scores (TES) Across Triage Days.
Note: Panel A shows the comparison of structured triage (NEXT and START-Neo-R) with untriaged evacuation. Panels B-F display the five representative patterns observed when comparing the quantitative (NEXT) and categorical (START-Neo-R) triage systems, ranging from complete separation to varying degrees of overlap. Abbreviations: TES, Total Evacuation Score; NEXT, Neonatal Extrication Triage; START-Neo-R, Simple Triage and Rapid Treatment for Neonates, Revised.

test how different prioritization rules behave before an actual disaster occurs. This may be particularly useful because NICU evacuation may be led by clinicians with limited disaster training or limited real evacuation experience.[5]

For facility-level planning, the TES framework may help NICUs compare alternative evacuation-order rules before a disaster occurs, identify how much prioritization depends on tied scores or categories, and make the assumptions behind evacuation plans

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 08:07:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1049023X26109157

explicit and testable. If externally validated, evacuation orders that reduce the burden of high-acuity patients remaining in the unit may plausibly reduce time pressure, staffing burden, equipment bottlenecks, or transport-resource constraints during evacuation. In this sense, the availability of a reproducible simulation metric is itself a practical methodological advance in a setting where controlled evaluation of actual evacuation is rarely feasible.

Research Implications

The broader implication of this work is methodological: TES can be extended from an equal-step model to a weighted model if empirical coefficients for transport time, staffing burden, equipment requirements, or transfer distance become available. For example, evacuation drills or transport simulations could estimate the additional time associated with mechanical ventilation, multiple infusion pumps, incubator transport, or hemodynamic instability; these coefficients could then be incorporated into a weighted TES model.

Future research should focus on multicenter datasets, standardized collection of variables required for multiple triage systems, empirical measurement of evacuation time and staffing burden, and linkage of simulated prioritization to real or drill-based outcomes. These steps would allow TES-based evaluation to progress from a proof-of-concept framework toward externally validated evacuation planning tools.

Limitations

This study has several limitations. First, it was conducted using a single-center dataset, and the patient case-mix, staffing model, equipment availability, transfer routes, and receiving capacity may not be representative of other institutions. Second, the source data were collected between 2019 and 2020. Changes in neonatal care, respiratory support, transport equipment, and NICU case-mix since that period may affect generalizability.

Third, this was a simulation-based study and did not evaluate actual evacuation time, adverse events during transport, mortality, deterioration after transfer, or receiving-unit outcomes. The TES should therefore be interpreted as an evacuation-order metric rather than a validated surrogate for clinical outcomes. Fourth, external validation of TES was beyond the scope of this study and should be addressed in future studies using real evacuation records, evacuation drills, transport-time measurements, or post-transfer outcome data. Fifth, NTISS was used as a reference severity measure because it had been recorded in the source clinical dataset and is validated as a therapy-based neonatal severity score, but it may not perfectly reflect logistical evacuation priority in every clinical scenario.

Sixth, the equal-step model does not capture variation in transport time, staffing burden, or equipment needs across patients. This simplification was used to isolate prioritization order, but future weighted models should incorporate empirical data on transport time, personnel requirements, transfer distance, and local resources. Seventh, additional triage systems were not applied retrospectively because the variables required to address them reliably were not recorded in a standardized manner, and post hoc classification could introduce misclassification bias. Finally, implementation during an actual disaster may be influenced by human factors, such as stress, fatigue, communication challenges, incident command decisions, or team performance; therefore, the observed TES advantage of NEXT should not be directly extrapolated to real-world evacuation effectiveness.

Conclusion

In this simulation study using real-world NICU data, structured triage generated more favorable evacuation-order distributions than evacuation without triage, and NEXT provided more consistent prioritization than START-Neo-R. The main contribution of this study is a reproducible TES-based framework for comparing neonatal evacuation-order strategies under explicit assumptions. Although these findings should not be interpreted as evidence that NEXT improves actual evacuation time or patient outcomes without further validation, the TES framework may help NICUs develop more transparent, testable, and adaptable evacuation plans.

Supplementary Materials. To view supplementary material for this article, please visit https://doi.org/10.1017/S1049023X26109157

Data Availability. The analysis dataset used for the evacuation simulations is provided as Supplementary Material 6.

Acknowledgements. The authors thank Ms. Takako Nishio and Ms. Yuko Karasawa for their valuable assistance with data collection, and all staff of the NICU at Nagoya City University Hospital for their continuous support throughout the study.

Author Contributions. **Kazunori Imai:** Conceived the study, designed the simulations, collected data, and drafted the manuscript.

Chunxi Xu: Contributed to data analysis and co-wrote parts of the methods section.

Shoya Oikawa: Conducted the simulations and performed the primary data analysis.

Yuko Araki: Supervised the statistical analysis.

Shinji Saitoh: Reviewed and provided critical feedback on the manuscript.

Osuke Iwata: Supervised the study design and overall project, and critically revised the manuscript.

All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Funding Statement. This study was supported by Grant-in-Aid for Scientific Research (C), grant number 21K11151.

Competing Interests. Kazunori Imai: The author declares none. Chunxi Xu: The author declares none. Shoya Oikawa: The author declares none. Yuko Araki: The author declares none. Shinji Saitoh: The author declares none. Osuke Iwata: The author declares none.

Use of AI Technology. Artificial intelligence (AI) tools, including ChatGPT (OpenAI - San Francisco, California USA; accessed during manuscript drafting and revision in 2025-2026) and Gemini (Google Inc. - Mountain View, California USA; accessed during manuscript revision in 2026), were used to assist with language editing, summarization, and drafting of figure legends. Model versions available through these platforms changed during the drafting and revision period; therefore, a single fixed model version was not specified. These tools were not used for data generation, statistical analysis, or scientific interpretation. All AI-assisted content was critically reviewed, verified, and revised by the authors.

References

1. Romanello M, Walawender M, Hsu SC, et al. The 2024 report of the Lancet Countdown on health and climate change: facing record-breaking threats from delayed action. *Lancet*. 2024;404(10465):1847–1896. doi: 10.1016/S0140-6736(24)01822-1.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 08:07:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1049023X26109157

2. Disaster Preparedness Advisory Council, Committee on Pediatric Emergency Medicine. Ensuring the health of children in disasters. *Pediatrics.* 2015;136(5): e1407–e1417. doi: 10.1542/peds.2015-3112.
3. Espiritu M, Patil U, Cruz H, et al. Evacuation of a neonatal intensive care unit in a disaster: lessons from Hurricane Sandy. *Pediatrics.* 2014;134(6): e1662–e1669. doi: 10.1542/peds.2014-0936.
4. Iwata O, Kawase A, Iwai M, et al. Evacuation of a tertiary neonatal center: lessons from the 2016 Kumamoto earthquakes. *Neonatology.* 2017;112(1): 92–96. doi: 10.1159/000466681.
5. Barfield WD, Krug SE, Committee on Fetus and Newborn, et al. Disaster preparedness in neonatal intensive care units. *Pediatrics.* 2017;139(5): e20170507. doi: 10.1542/peds.2017-0507.
6. Janvier A, Leblanc I, Barrington KJ. Nobody likes preemies: the relative value of patients' lives. *J Perinatol.* 2008;28(12):821–826. doi: 10.1038/jp.2008.103.
7. The Task Force of the Japanese Society for Prematurity, Neonatal Medicine and the Liaison Committee on Neonatal Care. Procedures for restoring the neonatal medical system in the event of a disaster 2020. https://jsnhd.or.jp/doctor/saigai/pdf/tejunsho_ver3_text.pdf. Accessed August 9, 2025.
8. Lin A, Taylor K, Cohen RS. Triage by resource allocation for IN-patients: a novel disaster triage tool for hospitalized pediatric patients. *Disaster Med Public Health Prep.* 2018;12(6):692–696. doi: 10.1017/dmp.2017.139.
9. Imai K, Suzuki T, Fukaya S, et al. A new quantitative triage system for hospitalized neonates to assist with decisions of hospital evacuation priorities. *Prehosp Disaster Med.* 2022;37(3):343–349. doi: 10.1017/S1049023X22000553.
10. Gray JE, Richardson DK, McCormick MC, et al. Neonatal therapeutic intervention scoring system: a therapy-based severity-of-illness index. *Pediatrics.* 1992;90(4):561–567. PMID: 1408510.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 08:07:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1049023X26109157