Financiado por la Unión Europea
NextGenerationEU

MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
SECRETARÍA DE ESTADO DE DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL

Plan de Recuperación, Transformación y Resiliencia

Engineering the future

Financiado por la Unión Europea NextGenerationEU
MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
SECRETARÍA DE ESTADO DE DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL
Plan de Recuperación, Transformación y Resiliencia

ARTIFICIAL INTELLIGENCE / BIOMEDICAL RESEARCH

Sepsis: identifying early signals in intensive care data

Early risk estimation service · Public technical note

Can we identify patterns of deterioration in a patient's condition before sepsis becomes apparent? EDGE/DM explores this question through artificial intelligence applied to historical intensive care unit (ICU) records.

This research capability examines how multiple measurements evolve over time and produces a signal to help investigate when signs of deterioration may emerge. It is not a diagnostic tool and is not intended to guide patient care.

40,336	1.55 M	2
ICU admissions	**hourly observations**	**hospital systems**
Public research dataset [2]	Longitudinal measurements	Diverse sources

The value of this Data Product

Demonstrating how health data can be turned into insights for research and made available as a service with controlled access, responsible use and traceability.

Research use only

The estimates reflect statistical patterns in the dataset studied. They must not be used to diagnose sepsis, trigger clinical alerts, select treatments or manage patients.

Explore the service in EDGE/DM: idneo.edgedm.eu/catalogue.html#sepsis-inference

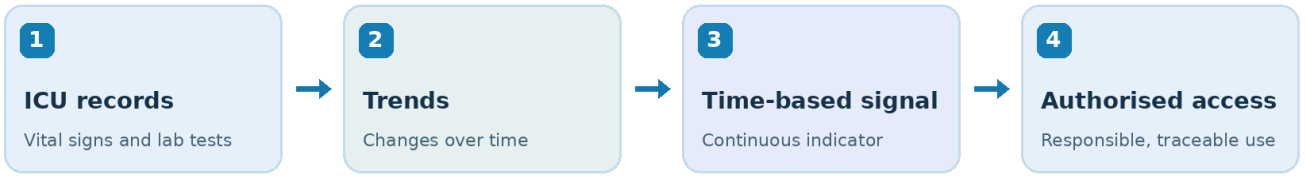
Financiado por la Unión Europea NextGenerationEU
MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
SECRETARÍA DE ESTADO DE DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL
Plan de Recuperación, Transformación y Resiliencia

1. From data to a time-based signal

The challenge is not to interpret a single measurement, but to understand how a patient's observed condition changes over time.

Sepsis is associated with a dysregulated response to infection that can cause severe organ dysfunction. Intensive care units record vital signs and laboratory results, but not all variables are measured equally often. This irregularity makes early signals harder to study.

How it works, in simple terms



The system considers both observed values and how they change over time. It also takes account of measurements that may be missing; an absent measurement is not automatically treated as a normal result.

Research dataset

Measure	Scope of the study
ICU admissions	40,336 admission records
Observations	1,552,210 hourly records
Admissions labelled with sepsis	2,932 admissions (7.27% of the total)
Data sources	Two hospital systems in PhysioNet/CinC 2019

The data come from the PhysioNet/Computing in Cardiology 2019 research challenge. Its labels follow an operational definition based on Sepsis-3. The figures describe a historical research cohort; they do not represent sepsis incidence across all hospitals.

How performance was evaluated

The model was evaluated on ICU admissions kept separate from those used for development. All hourly records from the same admission stayed together, preventing records from the same person being used in both development and evaluation. This was a retrospective assessment, not a prospective clinical study.

Financiado por la Unión Europea NextGenerationEU
MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
SECRETARÍA DE ESTADO DE DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL
Plan de Recuperación, Transformación y Resiliencia

2. Results: what the numbers really mean

Figures from an internal retrospective evaluation of the time-series model on test data from the two hospital systems.

How to read these results

Each measure answers a different question. These are not percentages of correct diagnoses and do not demonstrate clinical benefit. The metrics are shown together for transparency, but should not be added or averaged.

0.815

Ability to distinguish patterns

ROC-AUC · discrimination

When comparing a randomly selected hour labelled with sepsis against one without that label, the model assigns a higher score to the positive hour in approximately 82 out of 100 comparisons. This does not mean it correctly diagnoses 82% of patients.

0.412

How well the alert is timed

Utility Score · challenge-specific metric

This score rewards early warnings and penalises incorrect or late alerts. On the challenge's normalised scale, 0 represents a no-alert baseline and 1 the ideal outcome. A score of 0.412 does not mean 41.2% accuracy.

0.094

Detecting uncommon events

PR-AUC · precision–recall trade-off

This metric summarises how alert precision and the ability to find positive cases change as the decision threshold varies. It is particularly demanding when positive hours are rare. It does not mean 9.4% accuracy.

Responsible interpretation

These results indicate promising signals for research. However, they cannot tell us how many alerts would be correct in a real hospital or demonstrate patient benefit. No clinical performance has been established at a particular operational alert threshold.

Financiado por la Unión Europea NextGenerationEU
MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
SECRETARÍA DE ESTADO DE DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL
Plan de Recuperación, Transformación y Resiliencia

3. A research service within EDGE/DM

The service brings together advanced data analysis and responsible access.

Through the EDGE/DM catalogue, authorised organisations can discover the service and explore its use in research and technical evaluation projects. The aim is to support collaboration without making the underlying data or models freely downloadable.

Value for different stakeholders

Stakeholder group	Potential value
Researchers and universities	Explore methods for studying deterioration and complex clinical time series.
MedTech manufacturers and developers	Assess AI approaches relevant to future digital services.
Data governance and compliance teams	Examine an example of purpose-bound access, declared use and traceability.
Healthcare organisations and innovation teams	Identify possible research collaborations without assuming suitability for patient care.

Principles of responsible access

- **Defined purpose.** Use is limited to research, technical validation and non-clinical evaluation.
- **Data protection.** Direct patient identifiers are not needed to submit research queries.
- **Control of assets.** The service provides access to authorised results without requiring public release of source data.
- **Traceability.** Conditions of use and results can be linked to an authorised request.

Important limitations

> **Research is not clinical practice**
>
> Results are based on historical data and may differ with other populations, measurement practices or prevalence levels. No prospective clinical validation, independent evaluation in new hospitals or evidence of improved patient outcomes is presented. The generated signal must not be interpreted as a calibrated clinical probability.

Any move towards use in patient care would require specific studies of performance and safety, bias assessment, appropriate oversight and compliance with the relevant regulatory obligations.

Financiado por la Unión Europea NextGenerationEU
MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA
SECRETARÍA DE ESTADO DE DIGITALIZACIÓN E INTELIGENCIA ARTIFICIAL
Plan de Recuperación, Transformación y Resiliencia

4. Conclusions

An example of AI applied to health data, with an emphasis on responsible interpretation and controlled use.

Three key conclusions

- **Technical feasibility demonstrated.** Hourly ICU records can be analysed to identify statistical patterns associated with deterioration linked to sepsis. Internal testing achieved a ROC-AUC of 0.815.
- **Timing matters.** The analysis does more than classify observations: it also considers when the signal appears. The time-sensitive utility metric reached 0.412 under the research challenge rules.
- **Value beyond the algorithm.** EDGE/DM makes this capability available as a service with defined access conditions, purposes and traceability, opening opportunities for MedTech research and development collaboration.

What remains to be demonstrated

Before considering use in other settings, performance would need to be studied in new hospitals and populations, false-alert rates measured under representative conditions, and the system's safety, calibration and practical value validated. The current metrics do not resolve these questions.

Overall conclusion

The sepsis Data Product illustrates the potential of combining longitudinal health data, artificial intelligence and data governance. Its current contribution is technological and research-oriented: it enables the study of early deterioration signals and provides a documented starting point for further evaluation, without being presented as a clinical solution.

Public references

1. Reyna, M. A. et al. (2019). Early Prediction of Sepsis From Clinical Data: The PhysioNet/Computing in Cardiology Challenge. Critical Care Medicine. DOI: 10.1097/CCM.0000000000004145.

2. PhysioNet / Computing in Cardiology Challenge 2019. Dataset description and evaluation framework. https://physionet.org/content/challenge-2019/1.0.0/

3. Singer, M. et al. (2016). The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA. DOI: 10.1001/jama.2016.0287.