

Big data, AI and sepsis prediction: opportunities in the NICU

Daniel Vijlbrief · Neonatologist, Wilhelmina Children's Hospital, UMC Utrecht

CEPAS 2026 · Omics in Sepsis · Lyon, 31 October 2026

UMC Utrecht
Wilhelmina Kinderziekenhuis

Disclosure

Conflicts of interest

I have no conflicts of interest to declare

Artificial intelligence

I used AI (Claude) to prepare this presentation and the underlying research. Full disclosure at neonatology.ai.

Who used AI for poster or presentation preparation for this conference?

One very preterm infant, two months

~26 million

monitor values at one per second, five signals

rarely kept at this resolution

~5 GB

ECG waveform, one lead at 500 Hz

rarely kept

~90,000

words of clinical notes: a novel about one baby

kept in the record

>1 TB

continuous video, where it is recorded

few NICUs record it

What a sepsis model learns from

Events, not bytes

US cohort of 47,187 infants (22–28 weeks):

67% had at least one sepsis evaluation

10% of evaluations grew an organism

Order-of-magnitude estimates for a 60-day stay. Cohort: Speier et al., J Pediatric Infect Dis Soc 2023.

Heart rate and saturation are enough

½ to ⅔

of late-onset sepsis episodes flagged before clinical recognition

AUC 0.78–0.88 across groups, in different cohorts, outcomes and horizons: similar, not identical

Two signals suffice

Heart rate and SpO_2, one sample a minute. Hourly data loses the signal.

The algorithm barely matters

Logistic regression performs as well as gradient boosting or deep learning.

A pulse oximeter is enough

Pulse rate from the oximeter works as well as the ECG.

van den Berg 2023; Kausch 2023; Yang 2024; Meeus 2024. Seven groups mapped at neonatology.ai/sepsis.

Our own model: two numbers

≥47%

of infants with sepsis flagged before the blood culture

reported in the abstract

~4%

of alarms matched a true sepsis episode

reported in Table 3

Retrospective, one centre, 2008 to 2019

Blood-culture time stands in for first suspicion

False alarms should cost a look, not antibiotics

van den Berg M et al. Comput Biol Med 2023;163:107156. The model is openly available.

From data to better outcomes: four bridges

▼ the field is here



Bridge 1

Bytes to labelled episodes

Keep the data, link centres, date the first suspicion

Bridge 2

One unit to every unit

External and prospective validation; at the bedside a medical device (MDR)

Bridge 3

Alarm to the right decision

Three gates: look, treat, stop. The model fires at gate 1.

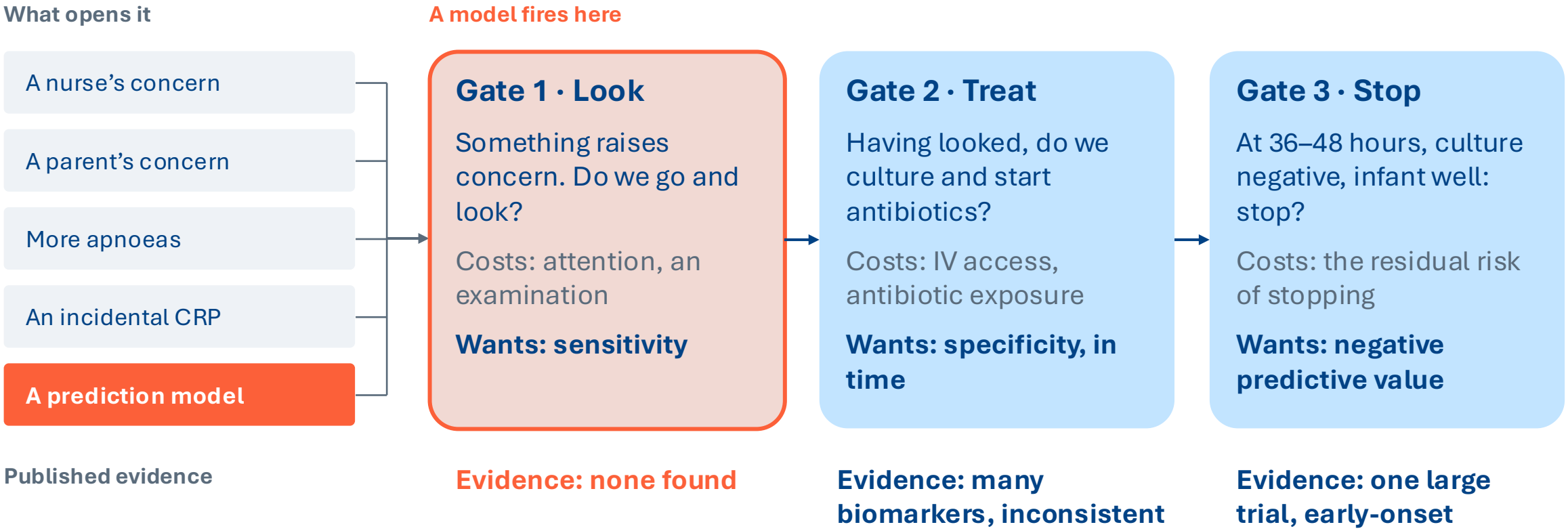
Bridge 4

Decision to proven benefit

One randomised trial in 15 years; test the response, not the model

Randomised trial: Moorman 2011.

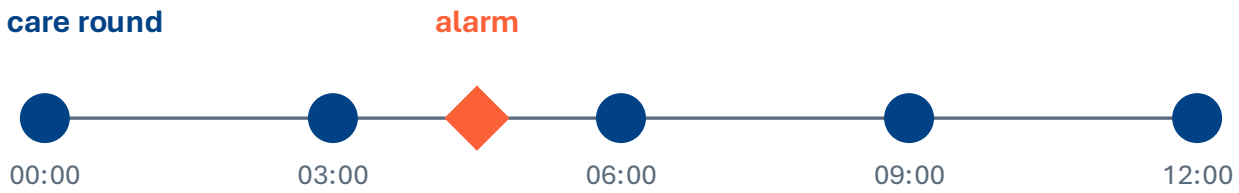
An alarm opens a cascade that already runs



Sola 2020; Stocker 2017; van Leeuwen 2024; Feng 2025. Gate 1: literature search, not a systematic review.

The model competes with the care round

Routine care already puts hands on every infant every few hours



Replace

Same number of looks; the model chooses who.

It has to choose better than the nurse.

Add

Extra looks on top of routine care.

They have to come early, in the right infants.

Does the model improve a look that would happen anyway, or create a useful look that otherwise would not?

Every model paper anchors onset to a recorded action, usually the blood culture. None counts the look before it.

Masino 2019; Cabrera-Quiros 2021; Mithal 2018; Mani 2014; van den Berg 2023.

Omics: each layer runs on its own clock

Layer	How fast it answers	Which decision it can reach
Monitor physiology, now		**Gate 1 · Look**
Gut microbiome, stool volatiles who is at risk this week		**A prior for gate 1**
Host RNA, metabolites only when blood is drawn		**Gate 2 · Treat**
Pathogen DNA PCR, cell-free DNA		**Gate 3 · Stop**
Genome mean nine days to report		**Mostly another question**



Graspeuntner 2019; Berkhout 2017; Thangavelu 2025; Oeser 2020; Sauberan 2026; Sloper 2024.

Big everything, small numbers

~780

transcripts per infant studied: 28,000 transcripts in 36 preterm infants

Cernada 2014

0.94 → 0.72

AUC of a four-gene birth signature, training versus external cohort

An 2024, term infants

0.79

five metabolites plus IL-6: sepsis versus non-infectious inflammation, at suspicion

Thangavelu 2025, 94 preterm infants

Not found

A model combining continuous physiology with a molecular layer for late-onset sepsis

literature search, not a systematic review

The opportunity

Decision support per gate, carrying the slower layers into the faster ones, combining the whole infant

More measurements per infant do not create more episodes. Multi-omics needs the same multicentre scale as physiology.

From one unit to the bedside

Clinical-impact trial
Does acting on the alarm change outcome?

Prospective silent validation
Does it hold in new infants, unseen?

Retrospective development
Would the alarm have fired earlier?

Nearly every model stops on the first rung

Generalisation
Few models validated across NICUs. Performance tends to fall in a new setting.

A medical device
Under the EU MDR (Rule 11) this software likely lands in Class IIb: notified-body review. Deployed is not validated: a widely used adult sepsis model reached AUC 0.63 externally.

Proof of benefit
One randomised trial in 15 years: 3,003 VLBW infants. Primary endpoint missed; mortality 10.2% to 8.1% as a secondary outcome.

Kausch 2023; Peng 2025; Wong 2021; Moorman 2011; Koppens 2023. MDR class: a reasoned reading, not a formal determination.

Share and collaborate

Open already
The model is available. Every search, paper and comparison behind this talk is online.

Possible soon
The European Health Data Space opens health data for research across borders.

Needed
More infants, more centres and more time than any one unit has.

No unit can build these bridges alone.

Let’s work together.

neonatology.ai/sepsis

the series, the literature base, the methods