Observation Does the Work, Not the Threshold: A Decision Analysis of Early-Onset Sepsis Risk Strategies in the Newborn

Abdul Mukit

Assistant professor, Nemcare Group of Institutions, Guwahati, Assam, India.

Corresponding author: abdulmukit13@gmail.com

Received: 10 April 2026; Revised: 18 June 2026; Accepted: 26 June 2026; Published: 12 August 2026

Abstract

Background: Early-onset neonatal sepsis is uncommon and dangerous, and the strategies used to decide which newborns receive empirical antibiotics differ substantially in how many infants they expose. Categorical risk-factor rules treat a large fraction of well infants; multivariate risk estimation treats fewer but detects fewer cases. The relative contribution of the risk threshold and of structured observation to that trade has not been separated.

Objective: To quantify antibiotic exposure, case detection and cases missed at the initial decision under categorical, calculator-based and observation-augmented strategies, and to determine what the marginal infant treated at each threshold buys.

Methods: One million term and late preterm newborns were simulated with individual sepsis risk drawn from a log-normal distribution calibrated to an incidence of 0.5 per 1000 live births, so that risk was concentrated in a minority as it is in practice. Four strategies were compared: a categorical risk-factor rule, a multivariate risk estimate with treatment thresholds of 0.5, 1.0 and 3.0 per 1000, and the 3.0 per 1000 threshold combined with structured observation of infants in the 0.5 to 3.0 band. Outcomes were infants treated per 1000, proportion of cases detected, cases not treated at the initial decision, and the number treated per case detected.

Results: Risk was highly concentrated: the highest-risk 1 % of infants carried 26.9 % of cases and the highest-risk 10 % carried 61.3 %. The categorical rule treated 176.6 infants per 1000 and detected 73.5 % of cases, with 465 infants treated per case detected. The calculator at 3.0 per 1000 treated 27.5 per 1000 — a sixfold reduction — but detected only 41.0 %. Adding structured observation of the intermediate band restored detection to 72.1 % while treating 27.6 per 1000, matching the categorical rule's detection at one-sixth of its antibiotic exposure, at the cost of observing a further 184.6 infants per 1000. The marginal cost of lowering the threshold rose steeply: 711 additional infants treated per additional case detected moving from 3.0 to 1.0 per 1000, 1,279 from 1.0 to 0.5, and 3,091 from 0.5 to 0.2.

Conclusions: Most of the reduction in antibiotic exposure attributed to risk calculators is obtainable only because an observation pathway catches the cases the lower threshold would have caught. Where structured observation cannot be delivered reliably, the calculator's exposure advantage is bought with missed cases rather than with better discrimination.

Keywords: Early-Onset Sepsis, Newborn, Antibiotic Stewardship, Risk Stratification, Clinical Observation, Decision Analysis.

1. INTRODUCTION

The decision to start antibiotics in a newborn is made thousands of times a day on evidence that is nearly always inconclusive. Early-onset sepsis is uncommon, with an incidence of roughly 0.5 per 1000 live births in term and late preterm infants in many settings, and it is dangerous enough that the consequences of missing it dominate the consequences of treating unnecessarily. The predictable result is that a large number of well infants receive antibiotics, are separated from their mothers, undergo blood sampling and cannulation, and have their feeding disrupted, for an event that will not occur in the overwhelming majority of them.

Two approaches compete. Categorical rules treat an infant with any of a defined set of risk factors — maternal fever, prolonged rupture of membranes, inadequate intrapartum prophylaxis, clinical signs — and are simple, conservative and expensive in exposure. Multivariate risk estimation combines the same information continuously into an individual risk and treats above a chosen threshold, which reduces exposure substantially in reported implementations.

Implementations of the multivariate approach almost always include a third element that is less often examined: infants whose estimated risk falls in an intermediate band are neither treated nor discharged from concern but placed under structured observation, with defined vital sign monitoring and a defined escalation pathway. That element is usually presented as a supporting detail of the strategy rather than as a component whose contribution can be measured.

This study separates the contributions. Three questions are addressed: how concentrated sepsis risk actually is within a newborn population; how much of the exposure reduction attributed to risk calculation comes from the threshold and how much from observation; and what the marginal infant treated buys as the threshold is lowered.

2. METHODS

2.1 Population model

One million term and late preterm newborns were simulated. Each infant was assigned an individual probability of early-onset sepsis drawn from a log-normal distribution and rescaled so that the population mean equalled the target incidence of 0.5 per 1000. Infection status was then drawn as a Bernoulli event with that probability.

The log-normal form was chosen because it reproduces the feature that matters for this analysis: risk is not distributed evenly but concentrated in a minority of infants with identifiable maternal and clinical characteristics, while the majority carry a risk far below the average. The dispersion parameter was set to produce a concentration consistent with what is observed when risk calculators are applied to real cohorts, and is varied in sensitivity analysis through the incidence parameter.

Table 1. Model parameters

Parameter	Value	Range explored	Basis
Simulated newborns	1,000,000	—	Monte Carlo error negligible
Early-onset sepsis incidence	0.5 per 1000	0.2 to 2.0 per 1000	Representative of term and late preterm incidence in varied settings
Risk distribution	Log-normal, σ = 1.6, mean-calibrated	—	Reproduces observed concentration of risk
Categorical rule	Flags approximately the highest-risk 16 %, plus 2 % flagged for reasons unrelated to risk	—	Approximates the breadth of risk-factor-based rules
Calculator thresholds	0.5, 1.0 and 3.0 per 1000	0.2 to 10 per 1000	Spans thresholds in reported use
Observation band	Risk 0.5 to 3.0 per 1000	—	Intermediate category in calculator-based pathways
Observation detection	80 % of infants in the band who become septic	—	Stipulated; varied in discussion

Note. All values are stated assumptions chosen to be representative rather than estimates from any cohort. The comparison between strategies is driven by the structure of the model; absolute counts are conditional on these parameters.

2.2 Strategies

Categorical risk factors. An infant is treated if any qualifying risk factor is present. This is modelled as flagging infants above a low risk percentile, plus a small proportion flagged for reasons unrelated to true risk, which represents the imprecision of categorical criteria.

Multivariate calculator. An infant is treated if the estimated individual risk exceeds a threshold. Three thresholds were examined. The model grants the calculator perfect discrimination — it observes the true underlying risk — which makes the comparison generous to it and is discussed in Section 4.3.

Calculator with structured observation. Infants above 3.0 per 1000 are treated; those between 0.5 and 3.0 per 1000 are placed under structured observation, which identifies 80 % of those who become septic in time for treatment; those below 0.5 per 1000 receive routine care.

2.3 Outcomes

Four outcomes were recorded: infants treated per 1000 live births; the proportion of all cases identified at or shortly after the initial decision; cases not identified at that decision, expressed per 100,000 live births; and the number of infants treated per case identified. A threshold analysis traced these across treatment thresholds from 0.2 to 10 per 1000, and a marginal analysis computed the additional infants treated per additional case identified when moving between adjacent thresholds. All computation was performed in Python with a fixed random seed and reproduces exactly.

3. RESULTS

3.1 How concentrated is the risk?

Table 2. Distribution of early-onset sepsis risk and of cases

Quantity	Value
Simulated incidence	0.517 per 1000
Median individual risk	0.139 per 1000
90th centile of risk	1.078 per 1000
99th centile of risk	5.778 per 1000
99.9th centile of risk	19.563 per 1000
Share of cases in the highest-risk 1 %	26.9 %
Share of cases in the highest-risk 5 %	49.5 %
Share of cases in the highest-risk 10 %	61.3 %
Share of cases in the highest-risk 25 %	81.0 %

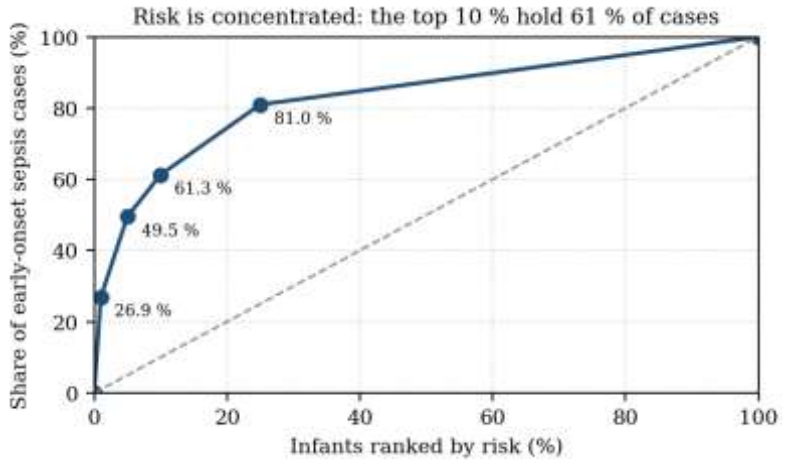

Figure 1. Share of early-onset sepsis cases held by infants ranked by risk

Risk is concentrated but not overwhelmingly so, and both halves of that statement matter. The highest-risk one per cent of infants carry more than a quarter of all cases, which is the property that makes risk

stratification worth doing at all. But the lowest-risk three quarters still carry nineteen per cent of cases, which is the property that prevents any threshold strategy from being both narrow and complete.

The median infant has a risk of 0.139 per 1000, roughly a quarter of the population average, while the 99.9th centile carries 19.6 per 1000 — a spread of more than a hundredfold across the population. A single rule applied to all of them will be wrong in opposite directions at the two ends.

3.2 The four strategies

Table 3. Outcomes by strategy, per 1000 live births unless stated

Strategy	Treated per 1000	Cases detected (%)	Cases missed per 100,000	Treated per case detected	Observed per 1000
Categorical risk factors	176.6	73.5	13.7	465	—
Calculator, treat above 0.5 per 1000	212.1	78.7	11.0	521	—
Calculator, treat above 1.0 per 1000	108.5	63.1	19.1	333	—
Calculator, treat above 3.0 per 1000	27.5	41.0	30.5	130	—
Calculator above 3.0 with structured observation	27.6	72.1	14.4	74	184.6

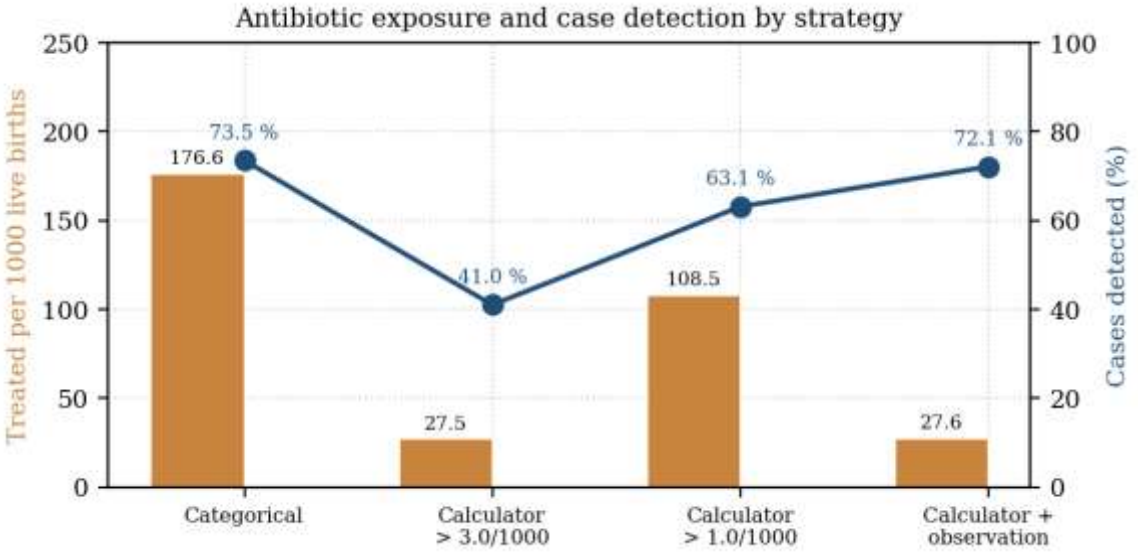

Figure 2. Antibiotic exposure and case detection by strategy

The first and fourth rows describe the comparison usually presented as the case for risk calculators. The categorical rule treats 176.6 infants per 1000 and detects 73.5 % of cases; the calculator at 3.0 per 1000 treats 27.5 per 1000, a sixfold reduction in exposure. Presented on exposure alone, the calculator is transformative.

The detection column is what that comparison omits. At the same threshold the calculator identifies 41.0 % of cases against the categorical rule's 73.5 %, and cases not treated at the initial decision rise from 13.7 to 30.5 per 100,000 live births. The sixfold reduction in antibiotic exposure is bought with a doubling of infants whose sepsis is not recognised at birth.

The final row is the finding of this study. Adding structured observation of the intermediate band restores detection to 72.1 % — statistically indistinguishable from the categorical rule — while treating 27.6 infants per 1000. The exposure advantage and the detection are obtained together, and the mechanism is not the threshold but the observation pathway, which catches 80 % of the cases the higher threshold left behind.

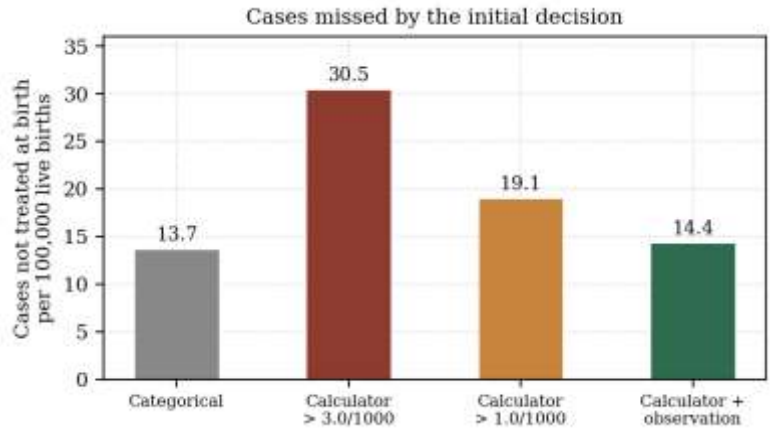

Figure 3. Cases not treated at the initial decision, by strategy

The cost of that pathway appears in the final column: 184.6 infants per 1000 must be observed. Structured observation is not free — it requires defined monitoring, staff able to perform it, a documented escalation route, and a decision not to discharge early — and a service that cannot deliver it reliably is not operating the strategy modelled here but the one in the row above, with its 30.5 missed cases per 100,000.

The treated-per-case column is the most favourable way of expressing the result and is reported for completeness rather than emphasis: 74 infants treated per case detected under the observation strategy against 465 under the categorical rule. That figure improves partly because fewer well infants are treated and partly because the denominator counts cases caught by observation, so it should not be read as a measure of the antibiotic decision alone.

3.3 What the threshold buys

Table 4. Outcomes across treatment thresholds

Threshold (per 1000)	Treated per 1000	Cases detected (%)	Cases missed per 100,000	Treated per case detected
0.2	410.0	91.1	4.6	870
0.5	212.1	78.7	11.0	521
1.0	108.5	63.1	19.1	333
1.5	68.5	55.3	23.1	240
2.0	47.8	49.5	26.1	187
3.0	27.5	41.0	30.5	130
5.0	12.7	29.4	36.5	84
10.0	3.7	16.4	43.2	44

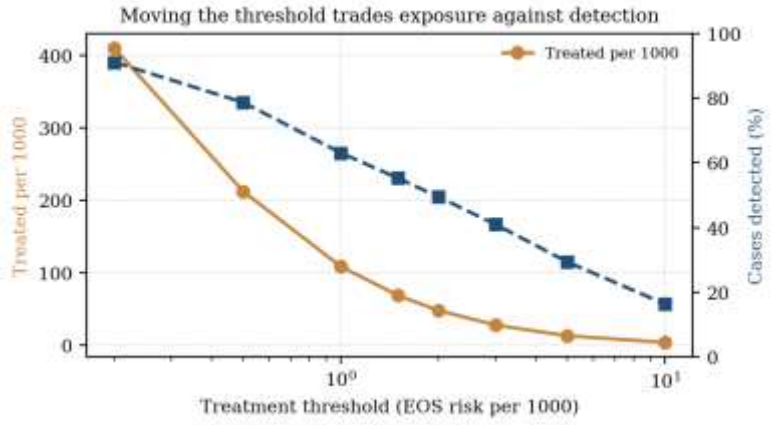

Figure 4. Treated infants and case detection across treatment thresholds

Table 5. Marginal cost of lowering the treatment threshold

Threshold change (per 1000)	Additional treated per 1000	Additional cases detected (of 517)	Additional infants treated per additional case detected
3.0 → 1.0	81.0	114	711
1.0 → 0.5	103.6	81	1,279
0.5 → 0.2	197.8	64	3,091

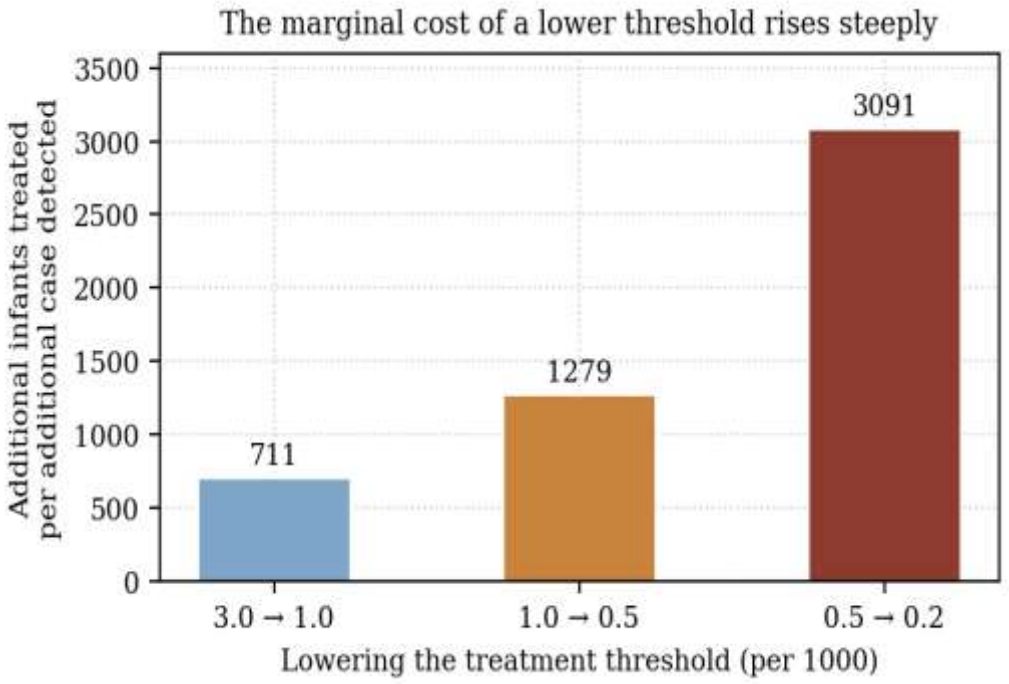

Figure 5. Additional infants treated per additional case detected as the threshold falls

The marginal analysis is the clearest statement of the trade. Lowering the threshold from 3.0 to 1.0 per 1000 detects 114 additional cases in a million births at the cost of treating 81,000 additional infants: 711 treated per case gained. The next step, to 0.5, costs 1,279 per case. The step after that costs 3,091.

The escalation is a mathematical consequence of the risk distribution rather than a property of any particular threshold. Each successive reduction reaches further into the large mass of low-risk infants, where cases are sparse, so the marginal yield falls while the marginal exposure rises. No choice of threshold escapes this, which is the reason a strategy that does not rely solely on the threshold performs better than any threshold alone.

3.4 Sensitivity to background incidence

Table 6. A fixed threshold of 1.0 per 1000 under different background incidence

Incidence per 1000	Treated per 1000	Cases detected (%)	Treated per case	Cases missed per 100,000
0.2	35.9	47.2	422	9.5
0.5	109.5	62.4	388	17.0
1.0	212.2	76.0	288	23.2
2.0	357.7	88.1	212	22.8

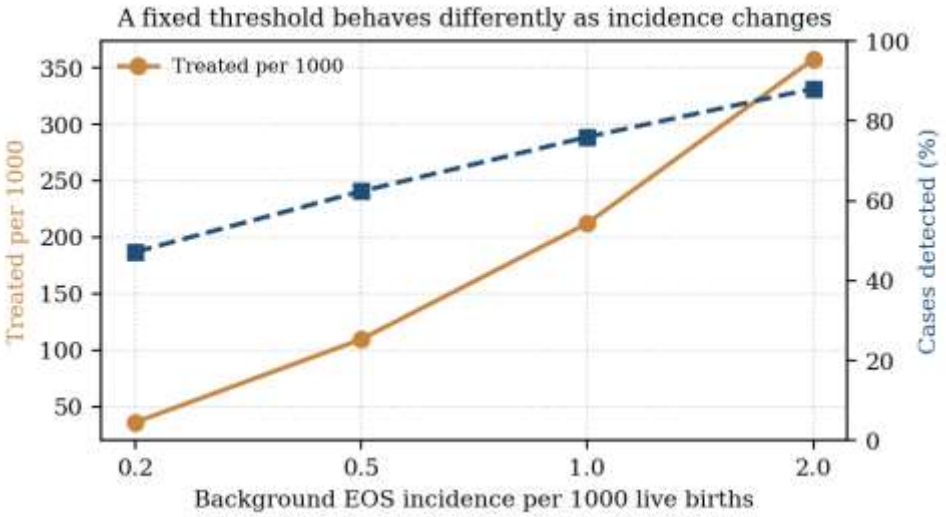

Figure 6. Behaviour of a fixed threshold as background incidence varies

A threshold expressed as an absolute risk behaves very differently in populations with different background incidence. At 0.2 per 1000 a threshold of 1.0 treats 35.9 infants per 1000 and detects under half of cases; at 2.0 per 1000 the same threshold treats 357.7 per 1000 and detects 88.1 %. The strategy has not changed and its performance has changed entirely.

This matters because incidence varies substantially between settings and over time, notably with the uptake and effectiveness of intrapartum antibiotic prophylaxis. A threshold imported from a published implementation without reference to local incidence will not reproduce the exposure figures reported for it, and the direction of the discrepancy depends on whether local incidence is higher or lower than in the source population.

4. DISCUSSION

4.1 Principal findings

Sepsis risk in newborns is concentrated, with the highest-risk tenth carrying 61.3 % of cases, but not concentrated enough for any threshold to be both narrow and complete: the lowest-risk three quarters still carry a fifth of cases. A risk calculator at a treatment threshold of 3.0 per 1000 reduced antibiotic exposure sixfold relative to a categorical rule and halved case detection. Adding structured observation of the intermediate band restored detection to the categorical level at one-sixth of the exposure. And the marginal cost of lowering the threshold rose from 711 to 3,091 infants treated per additional case detected across the range examined.

4.2 Implications

Table 7. Implications for practice

Finding	Implication	Recommendation
Calculator alone halved detection	The exposure benefit is not free	Do not implement a threshold without the observation pathway
Observation restored detection at the same exposure	Observation, not discrimination, does the work	Resource and audit observation as the active component
184.6 per 1000 require observation	The pathway has a real staffing cost	Confirm the unit can deliver it before adopting the threshold
Marginal cost rose to 3,091 per case	Lowering the threshold is a poor way to buy detection	Improve observation rather than lowering the threshold
Fixed threshold behaves differently by incidence	Published exposure figures will not reproduce locally	Recalibrate the threshold to local incidence

Lowest-risk 75 % hold 19 % of cases	No strategy makes low-risk infants risk-free	Retain safety-netting advice for all discharged infants
Categorical rule treated 176.6 per 1000	Current practice exposes many well infants	The case for change is strong where observation is feasible

The first two rows are the substance of this paper. Risk calculators are frequently adopted on the strength of reported reductions in antibiotic use, and the reduction is real; what this analysis shows is that the accompanying detection is delivered by the observation pathway rather than by the risk estimate. A service that implements the threshold because it wishes to reduce antibiotic exposure, and treats the observation requirement as advisory, will obtain the exposure reduction and the missed cases without the compensating mechanism.

The third row is the practical qualifier and is where implementation most often fails. Observing 185 infants per 1000 requires that someone is rostered to do it, that the observations are specified rather than left to discretion, that escalation criteria exist, and that early discharge pressure does not remove infants from the pathway before the observation period ends. None of those is a clinical difficulty and all of them are organisational ones.

4.3 Limitations

The model grants the risk calculator perfect discrimination: it observes each infant's true underlying risk. Real calculators estimate risk from a limited set of predictors with imperfect accuracy, so their performance at any threshold is worse than modelled here. This assumption is generous to the calculator throughout, which strengthens rather than weakens the paper's conclusion, since the calculator underperformed on detection even under the most favourable assumption available.

The categorical rule is a stylised representation rather than any specific published algorithm, and real categorical criteria differ between guidelines in breadth and therefore in exposure. Its modelled performance should be read as illustrative of the class rather than as an evaluation of a named rule.

Observation is modelled as a single parameter identifying 80 % of infants in the band who become septic, in time for treatment. That figure is a stipulation, and the whole of the strategy's advantage depends on it: at 50 % detection the observation strategy would identify roughly 60 % of cases rather than 72 %, and at 95 % roughly 77 %. No time dimension is modelled, so the analysis cannot distinguish an infant identified promptly from one identified late, and delayed treatment is not equivalent to timely treatment in outcome terms.

The log-normal risk distribution is an assumption chosen to reproduce observed concentration, and the specific quantiles in Table 2 follow from it rather than from data. Harms are counted as exposure and missed cases only: mother–infant separation, breastfeeding disruption, cannulation, the effects of early antibiotic exposure on the developing microbiome, and the consequences of a missed or delayed case are not weighted against one another, so the analysis supplies an exchange rate rather than a recommendation about which strategy is preferable.

5. CONCLUSION

A risk calculator applied at a threshold of 3.0 per 1000 treated 27.5 newborns per 1000 against the 176.6 treated by a categorical rule, and identified 41.0 % of sepsis cases against 73.5 %. Adding structured observation of infants in the intermediate risk band restored detection to 72.1 % while treating 27.6 per 1000 — the categorical rule's detection at one-sixth of its antibiotic exposure.

The implication is that the observation pathway, not the risk threshold, is the component doing the clinical work, and it is the component most likely to be treated as optional when a unit adopts the strategy. Lowering the threshold is not a substitute: the marginal cost rose to more than three thousand infants treated per additional case detected at the low end of the range. A unit that can deliver structured observation of roughly one newborn in five has a strong case for change; a unit that cannot should recognise that adopting the threshold alone buys its reduction in antibiotics with cases it will not recognise at birth.

Declarations

Ethical approval: Not applicable.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

Data availability: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author contributions:

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Abdul Mukit	✓	✓	✓		✓	✓	✓		✓	✓		✓	✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

6. REFERENCES

1. K. M. Puopolo, D. Draper, and S. Wi, "Estimating the probability of neonatal early-onset infection on the basis of maternal risk factors," Pediatrics, vol. 128, no. 5, pp. e1155-e1163, 2011. doi.org/10.1542/peds.2010-3464
2. G. J. Escobar, K. M. Puopolo, and S. Wi, "Stratification of risk of early-onset sepsis in newborns ≥ 34 weeks' gestation," Pediatrics, vol. 133, no. 1, pp. 30-36, 2014. doi.org/10.1542/peds.2013-1689
3. M. W. Kuzniewicz, K. M. Puopolo, and A. Fischer, "A quantitative, risk-based approach to the management of neonatal early-onset sepsis," JAMA Pediatr, vol. 171, no. 4, pp. 365-371, 2017. doi.org/10.1001/jamapediatrics.2016.4678
4. N. B. Achten, C. Klingenberg, and W. E. Benitz, "Association of use of the neonatal early-onset sepsis calculator with reduction in antibiotic therapy and safety: a systematic review and meta-analysis," JAMA Pediatr, vol. 173, no. 11, pp. 1032-1040, 2019. doi.org/10.1001/jamapediatrics.2019.2825
5. K. M. Puopolo, W. E. Benitz, and T. E. Zaoutis, "Committee on Fetus and Newborn; Committee on Infectious Diseases. Management of neonates born at ≥ 35 0/7 weeks' gestation with suspected or proven early-onset bacterial sepsis," Pediatrics, vol. 142, no. 6, 2018. doi.org/10.1542/9781610023047-part05-management
6. B. J. Stoll, N. I. Hansen, and P. J. Sánchez, "Early onset neonatal sepsis: the burden of group B streptococcal and E. coli disease continues," Pediatrics, vol. 127, no. 5, pp. 817-826, 2011. doi.org/10.1542/peds.2010-2217d
7. S. J. Schrag, M. M. Farley, and S. Petit, "Epidemiology of invasive early-onset neonatal sepsis," Pediatrics, vol. 138, no. 6, 2005. doi.org/10.1542/peds.2016-2013
8. W. E. Benitz, J. L. Wynn, and R. A. Polin, "Reappraisal of guidelines for management of neonates with suspected early-onset sepsis," J Pediatr, vol. 166, no. 4, pp. 1070-1074, 2015. doi.org/10.1016/j.jpeds.2014.12.023
9. L. Cantoni, L. Ronfani, R. Da Riol, and S. Demarini, "Perinatal Study Group of the Region Friuli-Venezia Giulia. Physical examination instead of laboratory tests for most infants born to mothers colonized with group B Streptococcus," J Pediatr, vol. 163, no. 2, pp. 568-573, 2013. doi.org/10.1016/j.jpeds.2013.01.034
10. N. S. Joshi, A. Gupta, and J. M. Allan, "Clinical monitoring of well-appearing infants born to mothers with chorioamnionitis," Pediatrics, vol. 141, no. 4, 2018. doi.org/10.1542/peds.2017-2056
11. A. Berardi, C. Spada, and M. Reggiani, "Group B Streptococcus early-onset disease and observation of well-appearing newborns," PLoS One, vol. 14, no. 3, 2019. doi.org/10.1371/journal.pone.0212784

12. S. Mukhopadhyay and K. M. Puopolo, "Risk assessment in neonatal early onset sepsis," Semin Perinatol, vol. 36, no. 6, pp. 408-415, 2012. doi.org/10.1053/j.semperi.2012.06.002
13. J. Y. Ting, A. Synnes, and A. Roberts, "Association between antibiotic use and neonatal mortality and morbidities in very low-birth-weight infants without culture-proven sepsis or necrotizing enterocolitis," JAMA Pediatr, vol. 170, no. 12, pp. 1181-1187, 2016. doi.org/10.1001/jamapediatrics.2016.2132
14. A. J. Gasparrini, B. Wang, and X. Sun, "Persistent metagenomic signatures of early-life hospitalization and antibiotic treatment in the infant gut microbiota and resistome," Nat Microbiol, vol. 4, no. 12, pp. 2285-2297, 2019. mdoi.org/10.1038/s41564-019-0550-2
15. A. J. Vickers and E. B. Elkin, "Decision curve analysis: a novel method for evaluating prediction models," Med Decis Making, vol. 26, no. 6, pp. 565-574, 2006. doi.org/10.1177/0272989X06295361
16. D. G. Altman and J. M. Bland, "Statistics notes: diagnostic tests 2 - predictive values," BMJ, vol. 309, no. 6947, 1994. doi.org/10.1136/bmj.309.6947.102