Averroes European Medical Journal
Open Access Publisher

November 2026 • Volume 1 • Number 1-Fr

Élaboration, validation et applicabilité clinique des modèles prédictifs du sepsis pédiatrique : revue systématique critique de la littérature (2021–2026)

Résumé

Contexte. Le sepsis pédiatrique demeure une cause majeure de mortalité évitable.
La révision consensuelle internationale de 2024 a écarté le syndrome de réponse inflammatoire systémique du champ diagnostique et retenu une définition fondée sur la dysfonction organique associée à une infection suspectée ou confirmée, traduite par le Phoenix Sepsis Score. Ce réajustement transfère l'enjeu clinique du constat de la défaillance organique vers son anticipation [3,4,5,6].
Objectif. Analyser de manière systématique les travaux portant sur l'élaboration et la validation de modèles prédictifs du sepsis de l'enfant et du nouveau-né, en caractérisant la capacité discriminative, la robustesse de la validation et la transposabilité clinique des instruments proposés.
Sources et sélection. Interrogation de PubMed/MEDLINE (janvier 2021 – novembre 2026). Ont été retenus les études de dérivation, les validations externes, les évaluations de dispositifs en déploiement clinique et les revues systématiques, méta-analyses et revues de portée.
Trente-quatre études ont été inclus. Les données ont été extraites selon une grille structurée : population source et niveau de soins, définition opérationnelle du sepsis, type de modèle employé et nombre de covariables, indices de discrimination et de calibration assortis de leurs intervalles de confiance, niveau de validation et biais identifiés.
Synthèse des résultats. Une méta-analyse en parapluie portant sur 18 revues systématiques rapporte une sensibilité poolée de 0,81 (IC 95 % 0,76–0,85) et une spécificité de 0,79 (IC 95 % 0,74–0,84), avec des aires sous la courbe comprises entre 0,78 et 0,94 selon le niveau de ressources [1]. Les études individuelles s'étendent de 0,56 à 0,99 [22]. Trois quarts des modèles néonatals (34 sur 44) n'ont fait l'objet que d'une validation unique [23]. L'évaluation d'un dispositif opérationnel, portant sur 166 943 passages aux urgences, révèle une sensibilité de 36 % (IC 95 % 29–43 %) et une valeur prédictive positive de 3 %, avec une sensibilité moins importante chez les enfants d'origine d'Afrique sub-saharienne, chez les enfants de moins de 5 ans et ne faisant pas partie de ceux pris en charge au sein des structures hospitalières pédiatriques tertiaires [2].
Conclusion. La dissociation entre les performances mesurées en recherche et celles observées en déploiement — de 94 % à 36 % de sensibilité — constitue le résultat structurant de la synthèse.
L'écart restreint séparant les algorithmes complexes des modèles parcimonieux indique que la performance prédictive dépend moins de la complexité algorithmique que de la qualité des covariables, de la répétition des validations externes et de l'équité de la performance entre sous-populations.

Mots-clés : sepsis pédiatrique ; choc septique ; détection précoce ; apprentissage supervisé ; Phoenix Sepsis Score ; validation externe ; équité en santé ; biomarqueurs endothéliaux.

Contributions des auteurs

Ahmed HADDADI[1] ; Karima HADDADI[2] ; Souad HADDADI[3]

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved

Averroes European Medical Journal Open Access Publisher
November 2026 • Volume 1 • Number 1-Fr

1. Introduction

Le sepsis est défini par une dysfonction organique menaçant le pronostic vital, secondaire à une réponse de l'hôte dérégulée par une infection. En pédiatrie, il demeure l'une des premières causes de mortalité mondiale, avec une prépondérance sur la période néonatale et dans les pays à revenus faibles ou intermédiaires [3,4].

La difficulté structurelle qui caractérise le sepsis est d'ordre diagnostique. Chez l'enfant le sepsis se manifeste initialement par une symptomatologie aspécifique — hyperthermie, tachycardie, polypnée, altération de l'état de conscience — qui se confond avec celle des infections non compliquées, mais dont l'évolution vers le choc septique peut survenir en quelques heures. Cette spécificité fait du retard diagnostique de facto le déterminant principal de la mortalité évitable.

Le syndrome de réponse inflammatoire systémique, utilisé comme critère diagnostique pendant trois décennies, a été formellement écarté en 2024 : sa sensibilité et sa spécificité insuffisantes le disqualifient pour poser le diagnostic de sepsis chez l'enfant [5]. Le consensus international retient désormais la notion de dysfonction organique associée à une infection suspectée ou confirmée, traduite par un Phoenix Sepsis Score supérieur ou égal à 2 points. Ce score composite intègre quatre domaines — cardiovasculaire, respiratoire, neurologique et hémostatique [6]. Il a été élaboré et validé sur plus de 3,6 millions de patients pédiatriques, hospitalisés dans 10 sites répartis sur 5 pays [6].

Cependant, le Phoenix Sepsis Score permet d'objectiver une dysfonction organique déjà constituée. Or, l'enjeu clinique réside dans l'anticipation de celle-ci. C'est ce déplacement — du constat vers la prédiction — qui structure la littérature des cinq dernières années et commande la lecture des résultats présentés ci-après dans cette étude.

2. Matériel et méthodes

2.1. Stratégie documentaire

Interrogation de PubMed/MEDLINE sur la période janvier 2021 – novembre 2026. Les stratégies de recherche ont associé, en champs titre-résumé et texte intégral, la définition du sepsis pédiatrique et néonatal (pediatric sepsis, neonatal sepsis, septic shock) à celles relatives à la prédiction et à la modélisation (prediction model, clinical prediction rule, machine learning, early warning score, external validation). Une recherche manuelle complémentaire a été conduite dans les bibliographies des travaux retenus.

2.2. Critères d'éligibilité

Critères d'inclusion. Études de dérivation de modèles prédictifs ; validations externes, y compris temporelles ; évaluations de dispositifs en déploiement clinique ; revues systématiques, méta-analyses et revues de portée portant sur la prédiction du sepsis de l'enfant ou du nouveau-né.

Critères de non-inclusion. Travaux dépourvus d'évaluation de performance prédictive ; études purement descriptives ou épidémiologiques ; publications sans indices de discrimination ou de calibration exploitables ; modèles dérivés et évalués exclusivement chez l'adulte.

2.3. Extraction des données et analyse critique

Les variables extraites concernent : la population source et le niveau de soins ; la définition opérationnelle du sepsis ; type de modèle employé et le nombre de covariables ; les indices de discrimination et de

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved

calibration assortis de leurs intervalles de confiance ; le niveau de validation (interne, externe, temporelle ou en déploiement) ; et les biais signalés par les auteurs ou identifiés lors de l'analyse critique.

Une grille d'évaluation structurée a été appliquée à chaque étude selon cinq axes : cohérence de la définition du critère de jugement avec les critères consensuels de 2024 ; prévalence de l'événement dans la population d'étude ; reproductibilité de la validation ; existence d'une analyse de sous-groupes portant sur l'équité ; démonstration d'un bénéfice clinique. Le risque de biais a été apprécié selon les domaines de l'outil PROBAST (Prediction model Risk Of Bias ASsessment Tool), qui distingue les domaines relatifs aux participants, au critère de jugement et à l'analyse [52].

3. Résultats

3.1. Épidémiologie : incidence, prévalence et déterminants microbiologiques

La caractérisation épidémiologique ne relève pas d'un préambule descriptif : elle détermine la valeur prédictive positive attendue d'un modèle, donc son utilité clinique effective.

Tableau 1. Indicateurs épidémiologiques du sepsis pédiatrique et néonatal.

Population	Indicateur	Valeur (IC 95 %)	Source
Enfants hospitalisés, États-Unis (2016–2023)	Incidence du sepsis	1,3 % (51 542/3 925 809)	[7]
	Mortalité hospitalière	10,1 %	[7]
	Sepsis d'origine communautaire	72,6 %	[7]
	Sepsis associé à un choc	61,6 %	[7]
	Sensibilité de la définition clinique	69,9 % (58,1–79,8)	[7]
	Spécificité de la définition clinique	93,1 % (89,6–95,7)	[7]
Enfants < 5 ans, Afrique	Prévalence poolée	40,2 % (26,8–55,4)	[8]
	Prévalence ajustée sur le biais de publication	27,1 % (15,5–43,0)	[8]
	Mortalité poolée	16,3 % (10,6–24,4)	[8]
Nouveau-nés, pays à revenus faibles	Prévalence poolée	40,98 % (30,50–51,46)	[9]
Prématurés extrêmes (22–28 semaines)	Incidence du sepsis précoce	3,3 % (4 354/130 938)	[10]
	Mortalité	35,5 %	[10]
	Mortalité ou morbidité	78,2 %	[10]
Nouveau-nés, unités de soins indiennes	Bactéries Gram négatif	70,0 % des isolats	[11]
	Multirésistance	75–88 % des isolats	[11]
	Létalité du sepsis documenté	36,6 % (12,1–71,0)	[11]

La lecture de ce tableau doit être conduite en tenant compte de deux éléments.
D'une part, les sensibilités et spécificités rapportées pour les définitions cliniques renseignent sur la performance de la définition elle-même et non sur celle d'un modèle prédictif.

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved

D'autre part, les prévalences issues de contextes à ressources limitées reflètent des populations hospitalières sélectionnées et ne sont pas transposables aux services d'urgence des pays à hauts revenus.

3.2. Performances discriminatives : synthèse des travaux 2021–2026

Le tableau 2 récapitule les travaux retenus. Il propose une juxtaposition de preuves hétérogènes et non une comparaison directe : les populations, les définitions opérationnelles du critère de jugement et les processus de validation diffèrent d'une étude à une autre, ce qui ne permet pas une comparaison quantitative des performances.

Tableau 2. Etudes retenues : population, type de modèle employé, performance et niveau de validation.

Étude	Population	Algorithme	Performance	Validation	Équité
Méta-analyse en parapluie, 2024 [1]	18 revues systématiques	Multiples	Sens. 0,81 ; Spéc. 0,79 ; AUC 0,78–0,94	Selon ressources	Non
Étude multicentrique, 2026 [13]	784 enfants, 2 centres	Forêt aléatoire, 12 covariables	AUC 0,887 (0,850–0,924) ; exactitude 82,1 %	Externe	Non
Modèle hématologique, 2026 [14]	Nouveau-nés et enfants	CatBoost, 18 covariables	ΔAUC +0,032 à +0,062 vs hémogramme	2 cohortes externes	Non
Régression LASSO, 2025 [15]	834 enfants	7 covariables	Discrimination et calibration satisfaisantes	Interne	Non
Biosenseurs, 2025 [16]	96 enfants, Dhaka	LASSO sur signaux physiologiques	AUC 0,78 (sepsis) ; 0,85 (choc) ; 0,87 (mortalité)	Validation croisée	Non
Modèle néonatal, 2026 [17]	Nouveau-nés, pays à revenus faibles et intermédiaires	Apprentissage supervisé	Sens. 90,3 % ; Spéc. 40,6 %	Interne	Non
Cohorte populationnelle, 2026 [18]	140 633 enfants, Hong Kong	Phoenix-8	AUC 0,91 (0,89–0,93)	Populationnelle	Oui
Validation saoudienne, 2026 [19]	431 admissions	Phoenix	Sens. 80,9 % ; VPP 27,0 %	Externe	Non
Modèle infantile, 2025 [20]	Enfants, réanimation	XGBoost	Sens. 0,77 ; Spéc. 0,74	Non précisée	Non
Étude princeps, 2025 [21]	1 604 422 passages	Gradient boosting	AUROC 0,94 (0,93–0,94)	Temporelle	Non
Dispositif commercialisé, 2026 [2]	166 943 passages	Dispositif intégré	Sens. 36 % ; VPP 3 %	Déploiemen	Oui — asymétrie

Premier axe — dispersion des performances et nature de ses déterminants.
Les aires sous la courbe s'étendent de 0,56 à 0,99 [22]. Une revue de portée portant sur 27 études relève que 85 % étaient monocentriques, 59 % recouraient à la régression logistique et 74 % utilisaient des données à faible prévalence de l'événement [22]. Cette dispersion procède donc principalement de divergences méthodologiques — effectifs, prévalence, critère de validation — et non de différences d'efficacité prédictive.
La conséquence est directe : une méta-analyse quantitative conduite sur ces données mesurerait l'hétérogénéité des méthodes plutôt que l'efficacité intrinsèque des algorithmes.

Deuxième axe — ampleur des gains liés à la complexité algorithmique.
La comparaison directe entre un modèle fondé sur 18 covariables hématologiques et un modèle fondé sur l'hémogramme seul objective des écarts d'aire sous la courbe de 0,032 à 0,062, avec une inversion dans l'une des cohortes externes [14].

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved

Une revue de portée rapporte par ailleurs une aire moyenne de 0,84 pour les modèles logistiques convertis en score, contre 0,81 pour les modèles linéaires et 0,85 pour les modèles arborescents [22]. Ces écarts, compris entre trois et six centièmes de point d'aire sous la courbe, sont du même ordre de grandeur que la variabilité inter-échantillons. Aucune conclusion quant à la supériorité d'une famille algorithmique ne peut en être tirée ; ils plaident en faveur de modèles parcimonieux, plus aisément interprétables par le clinicien et plus conformes aux exigences réglementaires de transparence.

Troisième axe — fragilité de la validation. Une revue de portée consacrée aux modèles néonatals dans les pays à revenus faibles et intermédiaires a recensé 82 études validant 44 modèles distincts chez 24 252 nouveau-nés. Trois quarts des modèles (34 sur 44) n'avaient fait l'objet que d'une validation unique ; deux tiers requéraient des paramètres biologiques ; quatre études seulement (5 %) étaient situées dans la région africaine sous supervision de l'Organisation mondiale de la Santé, une seule incluant des données issues d'un pays à faible revenu [23]. Cette configuration restreint la transposabilité des modèles précisément là où l'impact -incidence et prévalence- est le plus lourd.

Quatrième axe — déficit d'analyse d'équité. Sur les onze études du tableau, une seule a conduit une analyse de sous-groupes [2]. Pour autant cet état de fait ne constitue pas une lacune méthodologique accessoire : il prive cependant d'informations la performance d'un modèle produit les conséquences cliniques les plus graves.

Les quatre axes convergent vers un constat unique : la dispersion des performances rapportées mesure l'hétérogénéité des méthodes et des terrains, non une hiérarchie d'efficacité entre algorithmes.

3.3. Dissociation entre performance expérimentale et performance opérationnelle

L'évaluation multicentrique d'un dispositif opérationnel -commercialisé en l'espèce- de détection précoce, portant sur 166 943 passages aux urgences pédiatriques, mesure une sensibilité de 36 % (IC 95 % 29–43 %) et une valeur prédictive positive de 3 % (IC 95 % 2–4 %) pour l'identification d'un sepsis selon le Phoenix Sepsis Score, le déclenchement précédant le diagnostic dans 30 % des cas seulement [2]. La valeur prédictive négative dépasse 99 %.

Les analyses stratifiées constituent la donnée la plus préoccupante. La sensibilité était inférieure d'au moins dix points à la sensibilité globale chez les enfants d'origine africaine sub-saharienne (22 % ; IC 95 % 9–40 %), chez les enfants de moins de 5 ans (20 % ; IC 95 % 12–30 %) et dans les services d'urgence hors centres hospitaliers pédiatriques tertiaires (25 % ; IC 95 % 14–40 %) [2]. Les auteurs concluent que le dispositif « ne soutient pas un dépistage du sepsis pédiatrique opportun ou équitable ».

Quatre conclusions en découlent, qui doivent être formulées à certains égards sans atténuation.

Première conclusion. La dissociation entre 94 % et 36 % de sensibilité — de l'étude princeps au dispositif déployé — ne révèle aucun défaut méthodologique imputable aux auteurs de l'étude princeps [21]. Elle traduit le fait qu'un modèle testé sur des données structurées, complètes et régulièrement actualisées rencontrera, en routine, des données lacunaires et hétérogènes.
La performance prédictive n'est pas une propriété intrinsèque de l'algorithme : elle caractérise le couple algorithme–environnement informationnel.

Deuxième conclusion. Une valeur prédictive positive de 3 % associée à une valeur prédictive négative supérieure à 99 % définit un instrument de triage, non un instrument diagnostique. Sa finalité est d'exclure les patients chez qui le sepsis est hautement improbable. Un dispositif de triage évalué sur sa valeur prédictive positive apparaîtra toujours décevant ; évalué sur sa valeur prédictive négative, il peut présenter

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved

une utilité clinique réelle. Cette distinction engage la conception des systèmes d'alerte et la formation des équipes.

Troisième conclusion. L'asymétrie de performance selon l'origine ethnique, l'âge et le type d'établissement relève de l'équité en santé. Un modèle moins performant pour les populations les plus exposées produit l'inverse de l'effet recherché. Une revue de 2026 relève que les taux de faux positifs varient selon les niveaux de soins, révélant « des disparités critiques dans l'infrastructure des dossiers médicaux électroniques » [24].

Quatrième conclusion. Ces résultats n'invalident pas l'approche prédictive ; ils en définissent les conditions de validité. L'évaluation d'un modèle sur les données de l'établissement d'usage, assortie d'analyses de sous-groupes, constitue un prérequis au déploiement et ne fait pas double emploi avec l'évaluation conduite lors du développement.

3.4. Fondements physiopathologiques de la prédiction

Toute démarche prédictive postule qu'une phase de dérégulation détectable précède la défaillance organique. La littérature des cinq dernières années étaye ce postulat par cinq séries de travaux convergents.

3.4.1. Phénotypes microvasculaires identifiés par apprentissage non supervisé

Une étude portant sur 239 patients pédiatriques (âge médian 2,1 ans ; intervalle interquartile 0,5–11,1) a individualisé trois phénotypes microvasculaires distincts [25] :

Tableau 3. Phénotypes microvasculaires du sepsis pédiatrique et devenir associé.

Phénotype	Effectif	Caractéristiques microvasculaires	SDMV à 72 h	Mortalité
Préservé	52	Densité capillaire élevée, altération discrète du glycocalyx	10 %	1,9 %
Partiellement compensé	75	Activation inflammatoire marquée, réduction modérée de la densité capillaire	—	12,1 %
Décompensé	112	Densité capillaire minimale, défaillance microvasculaire globale	42 %	17,2 %

Le risque de syndrome de défaillance multiviscérale à 72 heures était maximal dans le phénotype décompensé comparé au préservé (rapport de risque instantané ajusté 2,47 ; IC 95 % 1,04–8,44 ; p = 0,042), indépendamment de l'âge, du foyer infectieux et du recours à la ventilation mécanique invasive. La mortalité suivait un gradient par paliers (1,9 %, 12,1 %, 17,2 % ; p = 0,023). Les analyses de validation interne ont reproduit l'attribution phénotypique et les associations pronostiques [25].

Cette étude apporte la validation empirique du postulat prédictif.
L'élément décisif réside dans l'indépendance : la microcirculation porte une information que les covariables cliniques et biologiques conventionnelles ne capturent pas.

3.4.2. Biomarqueurs du glycocalyx endothélial : mesure d'un mécanisme causal

Le mécanisme unissant l'activation inflammatoire à la défaillance organique est endothélial et quantifiable. Une étude portant sur 85 enfants septiques et 19 témoins appariés sur l'âge et le sexe établit que les concentrations sériques de syndécan-1, d'héparane sulfate et d'acide hyaluronique étaient significativement supérieures chez les patients septiques, et plus élevées dans les formes sévères que dans les formes légères ; l'acide hyaluronique était en outre majoré en présence d'un choc.
Dans la cohorte de choc septique, les non-survivants présentaient des concentrations supérieures à celles

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved

des survivants.
En analyse multivariée, le syndécan-1, l'acide hyaluronique et la procalcitonine étaient des prédicteurs indépendants de sepsis sévère.
Dans le choc septique, un syndécan-1 supérieur ou égal à 89,47 ng/mL prédisait une survie à 28 jours significativement réduite (p = 0,004) [26].

Une méta-analyse portant sur 17 études incluant 3 529 patients objective, chez les patients présentant un syndécan-1 élevé, un risque accru de décès (OR 2,32 ; IC 95 % 1,89–3,10), de défaillance multiviscérale (OR 3,3 ; IC 95 % 1,51–7,25) et d'insuffisance respiratoire (OR 7,53 ; IC 95 % 1,86–30,45) [27].

Trois propriétés confèrent à ces marqueurs une valeur prédictive : ils quantifient un mécanisme causal et non un marqueur non spécifique de gravité ; leur cinétique précède la défaillance cliniquement objectivée ; le syndécan-1 est dosable en circulation, ce qui autorise son intégration dans un modèle multimodal.
Leur limite actuelle réside dans l'absence de validation prospective au sein d'un modèle composite.

3.4.3. Endotypes immunologiques et cibles thérapeutiques identifiables

Une étude intégrant phénotypage immunitaire approfondi, protéomique plasmatique, transcriptomique unicellulaire et cytométrie de phosphoflux chez 88 enfants en état critique a individualisé trois sous-groupes immunologiques, dont un groupe de sévérité élevée caractérisé par une hypercytokinémie dépendante de l'IL-6 et de l'IFN-γ.
Ce groupe présentait une activation lymphocytaire T avec suppression d'un programme génique lymphoïde protecteur dans les sous-populations CD8+.
La cytométrie de phosphoflux objectivait une hyperactivation basale de STAT1/STAT3 dans ces lymphocytes, lesquels ne répondaient plus à la stimulation [28].

L'axe JAK/STAT est ainsi identifié comme cible thérapeutique rationnelle [28].
Ces travaux ouvrent une perspective précise : la prédiction pourrait porter non sur la survenue du sepsis en général, mais sur une trajectoire physiopathologique déterminée, permettant une intervention ciblée sur le mécanisme dominant.

Une étude complémentaire a individualisé deux phénotypes de choc septique pédiatrique par analyse de profil latent, associés à des signatures transcriptomiques de dysfonction endothéliale microvasculaire et de renouvellement neutrophile accru ; ces phénotypes ne présentaient pas de chevauchement avec les endotypes établis sur l'expression génique et ne montraient pas de réponse différentielle aux corticoïdes [29].
Ce résultat indique que les classifications phénotypiques et endotypiques ne se superposent pas, et qu'une prédiction fondée sur l'une pourrait ne pas restituer l'information portée par l'autre.

3.4.4. Cinétique des biomarqueurs plasmatiques et panels protéomiques

Une étude consacrée aux biomarqueurs dynamiques établit que les valeurs initiales et les décroissances précoces de la protéine C réactive, de la procalcitonine, de l'interleukine-6 et des D-dimères reflètent la dysfonction organique et la durée d'hospitalisation.
Les profils inflammatoires initiaux différaient selon l'étiologie — protéine C réactive et procalcitonine approximativement doublées dans les infections bactériennes comparativement aux infections virales, interleukine-6 maximale dans les infections respiratoires et abdominales.
Tous les marqueurs diminuaient significativement sur sept jours, avec des décroissances proportionnelles uniformes quel que soit le pathogène.

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved

Un algorithme en deux étapes — protéine C réactive initiale supérieure ou égale à 60 mg/L, procalcitonine supérieure ou égale à 3 ng/mL, interleukine-6 supérieure ou égale à 200 pg/mL ou D-dimères supérieurs ou égaux à 1,5 mg/L, associés à une décroissance d'au moins 50 % de l'interleukine-6 ou de la procalcitonine en 72 heures — identifiait les enfants nécessitant une prise en charge en réanimation ou une hospitalisation d'au moins sept jours avec une aire sous la courbe de 0,91 [30].
L'intérêt de ce résultat réside dans l'accessibilité des marqueurs, disponibles dans la majorité des laboratoires, y compris en contexte de ressources contraintes, par opposition aux technologies de recherche mobilisées par les approches multi-omiques.

Une protéomique sérique néonatale a par ailleurs identifié un panel protéique relevant de trois processus biologiques : détection du pathogène (LBP, CD14) et inflammation de phase aiguë (CRP, SAA1/2, ORM1/2) ; activation immunitaire innée et recrutement leucocytaire (FCGR3A, CSF1R, CD163, CD206) avec épuisement plaquettaire (PF4, PPBP, THBS1, GP5) ; et lésion endothéliale avec dysfonction microvasculaire (ICAM1, VCAM1, VWF, SPARC), accompagnée d'une perte simultanée de lipoprotéines et de serpines protectrices [31].

3.4.5. Immaturité immunitaire et cinétique de la défaillance

Un déterminant conditionne la faisabilité même de la prédiction en pédiatrie et doit être explicité.
Comparé à l'adulte, le nouveau-né présente des réponses médiées par les récepteurs de reconnaissance de motifs moléculaires (Toll-like receptors) atténuées et une production réduite de cytokines pro-inflammatoires ; l'expression de TLR4 à la surface des polynucléaires neutrophiles néonatals est basse, de même que celle des transducteurs de signal en aval, ce qui limite les réponses cytokiniques et la chimiotaxie [32].

Cette immaturité ne se traduit toutefois pas par une hypo-réactivité.
Les macrophages pédiatriques produisent davantage de TNF-α pro-inflammatoire et davantage d'IL-10 anti-inflammatoire que les macrophages adultes ; dans un modèle ex vivo de macrophages d'enfant stimulés par lipopolysaccharide, le rapport IL-10/TNF-α était quinze fois supérieur à celui observé chez l'adulte [32,33].
La traduction clinique est directe : l'enfant peut développer une hyper-inflammation et une immunosuppression concomitante au cours des 72 premières heures d'hospitalisation, avec une prévalence de dysfonction immunitaire simultanée supérieure à celle des cohortes adultes.
Cette trajectoire chevauchante, associée à l'installation simultanée plutôt que séquentielle du syndrome de défaillance multiviscérale, explique une mortalité précoce significativement supérieure à celle de l'adulte [34].
Conséquence pour la modélisation prédictive.
La survenue plus précoce de la défaillance, simultanée dans plusieurs systèmes, restreint la fenêtre d'anticipation et atténue la progressivité du signal d'alerte.
Cette contrainte biologique explique également pourquoi un modèle validé chez l'adulte ne peut être transposé à l'enfant sans validation propre.

3.5. Analyse critique des limites méthodologiques

3.5.1. Hétérogénéité des définitions opérationnelles

La revue de portée consacrée aux technologies prédictives relève que « les variations dans les définitions des critères de jugement soulignent l'importance potentielle des critères de consensus de 2024 dans le

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved

développement futur » [22]. Un modèle construit et testé sur des critères antérieurs mesure la survenue d'une entité clinique qui n'est plus reconnue. L'alignement sur les critères Phoenix ne relève donc pas d'un choix méthodologique parmi d'autres : il constitue une condition de validité.
L'hétérogénéité temporelle introduite par la refonte de 2024 interdit de comparer directement les performances des modèles antérieurs et postérieurs.

3.5.2. Faible prévalence et conséquences cliniques

Sur les 27 études de la revue de portée, 74 % utilisaient des séries de données à faible prévalence [22].
La conséquence est mécanique : un modèle dont la sensibilité et la spécificité sont élevées, appliqué à une population où l'événement est rare, génère une proportion importante de fausses alertes — ce que confirme la valeur prédictive positive de 3 % mesurée en déploiement [2].
Cet effet comporte une conséquence secondaire souvent négligée : la multiplication des fausses alertes induit une fatigue d'alerte au sein des équipes soignantes, réduisant progressivement la réactivité aux alertes pertinentes.
Performance statistique et effet clinique ne sont donc pas seulement distincts : ils peuvent évoluer en sens inverse.

3.5.3. Équité : limite la plus sévère

L'asymétrie de sensibilité observée chez les enfants d'origine africaine sub-saharienne, les moins de 5 ans et hors centres tertiaires [2] ne procède pas d'un défaut algorithmique isolé : elle reflète l'inégale qualité de la documentation dans les systèmes d'information et la composition différente des populations selon les établissements.
Un modèle moins performant pour les populations les plus exposées produit l'inverse de l'effet recherché.
Sur les onze travaux analysés, un seul conduit une analyse d'équité.
Cette absence ne relève pas d'une négligence individuelle des auteurs, mais d'une norme de publication qui n'a pas encore intégré l'équité comme critère de validation.

3.5.4. Validité du critère de jugement

Une revue systématique critique relève que la mortalité hospitalière toutes causes, utilisée comme critère de jugement dans la cohorte de développement des critères Phoenix, « peut ne pas refléter pleinement la mortalité associée au sepsis et pourrait introduire une classification erronée des résultats » [35].
Un critère de jugement mesurant la mortalité toutes causes attribuera au sepsis des décès relevant d'autres déterminants, et inversement.

3.5.5. Limites méthodologiques explicitées

Plusieurs limites doivent être annoncées d'emblée, car elles conditionnent la portée des conclusions.
Les performances rapportées proviennent de travaux dont les populations, les définitions opérationnelles et les algorithmes diffèrent, ce qui proscrit toute agrégation quantitative — proscription qui constitue en elle-même un résultat, comme le développement l'établira.
Il est à noter également que la période couvre 2024, année de refonte des critères diagnostiques, ce qui introduit une hétérogénéité temporelle entre les travaux antérieurs et postérieurs. Enfin, les travaux négatifs ou dépourvus de performance exploitable étant moins susceptibles d'être publiés, un biais de publication non quantifiable subsiste.

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved

4. Discussion

4.1. Acquis des cinq dernières années

Premier acquis. La détection précoce du sepsis pédiatrique repose sur un fondement biologique établi. Les phénotypes microvasculaires précèdent et prédisent la défaillance indépendamment des covariables conventionnelles [25] ; les biomarqueurs du glycocalyx quantifient un mécanisme causal dont la cinétique précède la défaillance clinique [26,27] ; les profils protéomiques associent détection du pathogène, activation immunitaire et lésion endothéliale [31].
Ces résultats, issus de populations distinctes et de méthodes distinctes, convergent : la phase de dérégulation est détectable avant la défaillance.

Deuxième acquis. La performance prédictive dépend moins de l'amelioration algorithmique que de la qualité des covariables et de la robustesse de la validation.
Les écarts entre modèles complexes et modèles parcimonieux sont du même ordre de grandeur que la variabilité inter-échantillons [14,21,22], et trois quarts des modèles néonatals n'ont fait l'objet que d'une validation unique [23]. L'amélioration computationnelle ne constitue donc pas le levier d'amélioration prioritaire ; la qualité des données et la répétition des validations le sont.

Troisième acquis. L'implémentation constitue le goulet d'étranglement, non la modélisation.
La dissociation entre 94 % et 36 % de sensibilité [2,21] mesure la distance séparant la démonstration statistique de l'effet clinique.
Un modèle dont la performance n'est validée que sur des données de recherche n'a pas établi sa capacité à modifier le devenir des patients.

4.2. Articulation entre prédiction et intervention thérapeutique

La prédiction n'a de valeur que si l'intervention déclenchée est efficace. Les essais récents rappellent que cette efficacité n'est ni universelle ni intuitive — rappel qui constitue une contrainte majeure pour le champ.

Les recommandations 2026 de la Surviving Sepsis Campaign pédiatrique ont révisé substantiellement la réanimation volémique : administration suggérée de 40 à 60 mL/kg en bolus fractionnés (10 à 20 mL/kg par bolus) au cours de la première heure en contexte de disponibilité de réanimation (recommandation conditionnelle, faible niveau de preuve) ; abstention formellement recommandée chez l'enfant présentant un sepsis sans hypotension en contexte d'absence de réanimation (recommandation forte, niveau de preuve élevé) [36].

L'essai PRoMPT BOLUS, portant sur 9 041 enfants traités pour choc septique, ne retrouve pas de différence significative entre cristalloïdes balancés et sérum salé isotonique sur le critère principal (3,4 % contre 3,0 % ; différence 0,4 point ; IC 95 % −0,5 à 1,3 ; risque relatif 1,10 ; IC 95 % 0,88–1,40), malgré une réduction de l'hyperchlorémie (31,4 % contre 49,0 %) [37].
Le nombre médian de jours sans hospitalisation à 28 jours était identique (23 jours).

Une méta-analyse actualisée portant sur neuf études confirme l'absence de réduction de la mortalité (RR 0,94 ; IC 95 % 0,83–1,06 ; $I^2 = 0$ %) et de l'insuffisance rénale aiguë (RR 0,84 ; IC 95 % 0,61–1,16 ; $I^2 = 73,5$ %), la taille d'information requise n'étant pas atteinte, ce qui laisse ces résultats non détermients.
Elle objective en revanche une réduction du recours à l'épuration extrarénale (RR 0,74 ; IC 95 % 0,56–0,96 ; $I^2 = 23,5$ %) et de l'hyperchlorémie (RR 0,65 ; IC 95 % 0,61–0,69 ; $I^2 = 0$ %) [38].

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved

Concernant les vasopresseurs, une étude de cohorte appariée par score de propension établit que l'administration de noradrénaline était associée à une mortalité inférieure à celle de l'adrénaline chez les enfants sans dysfonction cardiaque connue (3,7 % contre 0 % ; différence de risque 3,7 % ; IC 95 % 0,2–7,2) ; il s'agit d'une association observationnelle, exposée au biais de confusion résiduelle par indication, et non d'une relation causale établie [39].
Une étude observationnelle prospective multicentrique portant sur 225 patients relève que l'adrénaline demeurait l'agent de première intention dans 70 % des cas, suivie de la noradrénaline dans 27 % [40].

Concernant la corticothérapie, les recommandations 2026 suggèrent de ne pas recourir à l'hydrocortisone intraveineuse lorsque la stabilité hémodynamique peut être rétablie par la réanimation volémique et les vasoactifs (recommandation conditionnelle, faible niveau de preuve), et estiment les données insuffisantes pour formuler une recommandation dans les chocs réfractaires [36].

Ces résultats délimitent le cadre dans lequel une alerte prédictive produit son effet — cadre étroit. Anticiper l'identification d'un patient ne modifie son pronostic -devenir- qu'à la condition que l'intervention qui sera instaurée soit efficace.
C'est la raison pour laquelle la démonstration d'un effet clinique, et non la performance statistique isolée, constitue le critère ultime de validité d'un modèle.

4.3. Conditions de validité d'un modèle déployable

La synthèse permet d'énoncer six conditions de validité, hiérarchisées selon leur caractère nécessaire et leur faisabilité opérationnelle.

Première — Alignement sur les critères consensuels de 2024. Un modèle testé sur les critères antérieurs mesure une entité clinique qui n'est plus reconnue [5,6].

Deuxième — Validation externe, et si possible répétée. Trois quarts des modèles néonatals n'ont fait l'objet que d'une validation unique [23] ; un modèle validé sur sa population d'origine n'a établi que sa capacité d'ajustement.

Troisième — Validation temporelle. Un modèle testé sur des données antérieures à la pandémie peut voir sa performance modifiée par l'évolution de l'épidémiologie, des pratiques et de la documentation.

Quatrième — Analyses stratifiées d'équité. L'asymétrie de performance selon l'origine ethnique, l'âge et le type d'établissement [2] impose que l'équité constitue un critère de validation et non une optimisation optionnelle.

Cinquième — Évaluation sur les données de la vraie vie. La dissociation entre recherche et déploiement [2,21] établit que cette évaluation ne fait pas double emploi avec celle conduite lors du développement.

Sixième — Démonstration d'un bénéfice clinique. Aucun des travaux analysés n'a conduit d'essai randomisé comparant une décision assistée par modèle à une décision clinique conventionnelle ; plusieurs auteurs reconnaissent explicitement que cette évaluation reste à conduire [17,22]. Cette insuffisance centrale de la littérature actuelle peut invalider l'interprétation des performances rapportées comme preuves d'efficacité.

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved

4.4. Perspectives

Trois orientations paraissent prioritaires, par ordre de faisabilité décroissante.

Première — Prise en compte des biomarqueurs liés aux mécanismes physiopathologiques. Les biomarqueurs du glycocalyx quantifient un mécanisme causal [26,27] et le panel protéomique néonatal associe détection du pathogène, recrutement leucocytaire et lésion endothéliale [31]. Leur intégration pourrait accroître la spécificité, là où les modèles fondés sur les seules covariables conventionnelles plafonnent — sous réserve d'une validation prospective au sein d'un modèle multimodal, non encore conduite.

Deuxième — Stratification par endotype. L'individualisation d'un endotype ciblable associant signalisation STAT aberrante et dérégulation lymphocytaire T [28] ouvre la voie à une prédiction portant non sur la survenue du sepsis, mais sur une trajectoire physiopathologique déterminée. Cette orientation demeure conditionnée par la production de données multi-omiques à grande échelle, ce qui en limite la faisabilité immédiate.

Troisième — Évaluation de l'impact clinique, priorité absolue. Une revue relève qu'aucun travail n'a évalué l'impact clinique des technologies prédictives du sepsis pédiatrique [22]. En l'absence d'essai randomisé comparant une décision assistée par modèle à une décision clinique conventionnelle, l'ensemble des performances rapportées mesure une capacité prédictive et non un bénéfice pour le patient.

4.5. Limites de la présente revue

Cette synthèse repose sur une recherche est restreinte à la période 2021–2026. Aucun instrument formel d'évaluation quantitative du risque de biais n'a été appliqué, et les performances rapportées proviennent de travaux dont les populations, les définitions opérationnelles et les algorithmes diffèrent. Cette hétérogénéité, loin de constituer un simple défaut méthodologique, représente en elle-même un résultat : elle établit qu'une méta-analyse quantitative conduite sur ces données mesurerait l'hétérogénéité des méthodes plutôt que l'efficacité intrinsèque des modèles.

5. Conclusion

Trois conclusions se dégagent, par ordre de solidité décroissante.

Première conclusion. La détection précoce du sepsis pédiatrique repose sur un fondement biologique établi. Les phénotypes microvasculaires identifiés par approche sans a priori précèdent et prédisent la défaillance d'organe indépendamment des variables cliniques conventionnelles, avec une mortalité variant de 1,9 % à 17,2 % selon le profil observé.

Parallèlement, les biomarqueurs du glycocalyx permettent de quantifier un mécanisme physiopathologique causal directement corrélé au pronostic.

Chez l'enfant, la singularité de la réponse immunitaire — marquée par un rapport IL-10/TNF-α quinze fois supérieur à celui de l'adulte et par la coexistence d'une hyper-inflammation et d'une immunosuppression — rend cette fenêtre d'anticipation à la fois indispensable et particulièrement étroite [32,33,34].

Deuxième conclusion. La performance prédictive dépend moins de la complexité de la modélisation que de la qualité des covariables et de la robustesse de la validation.

Les écarts entre modèles complexes et modèles parcimonieux sont du même ordre de grandeur que la

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved

variabilité inter-échantillons [14,21,22], et trois quarts des modèles néonatals n'ont fait l'objet que d'une validation unique [23].

Troisième conclusion. Le principal facteur limitant réside désormais dans l'intégration clinique de ces outils, et non plus dans leur modélisation. L'effondrement de la sensibilité, chutant de 94 % dans l'étude princeps à 36 % en pratique courante [2,21] — avec une perte de performance particulièrement marquée chez les enfants d'origine africaine sub-saharienne, les moins de 5 ans et en dehors des centres tertiaires — engage directement la responsabilité des concepteurs comme des autorités de déploiement.

La lacune majeure demeure l'absence d'évaluation de l'impact clinique.

Aucun essai randomisé n'a comparé une décision fondée sur un modèle prédictif à une décision clinique conventionnelle [17,22]. Tant que cette démonstration n'est pas produite, la valeur clinique des modèles prédictifs du sepsis pédiatrique demeure une hypothèse biologiquement fondée et méthodologiquement non établie.

Déclarations

Contributions des auteurs (CRediT) : Ahmed HADDADI[1] ; Karima HADDADI[2] ; Souad HADDADI[3]
[1.] MD PhD, Polyclinique Medical GROUPE AVERROES; Research Unit EA 2694 Public Health, Epidemiology, and Quality of Care, Medical University, Lille, France a.haddadi@averroes-emj.com
[2.] MD, Msc, Hopital Antoine Becler ; karima.haddadi@aphp.fr
[3.] PhD, Polyclinique Medical groupe AVERROES ; souad.h@averroes-emj.com
Conflits d'intérêts : les auteurs déclarent n'avoir aucun conflit d'intérêts en lien avec ce travail.
Financement : ce travail n'a bénéficié d'aucun financement.
Tous les auteurs ont approuvé la version soumise.
Approbation éthique : revue narrative sans recueil de données individuelles ; aucune approbation éthique n'était requise. Les données analysées proviennent exclusivement de la littérature publiée

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved

Références

1. Umbrella review for pediatric sepsis and robust prediction potential worldwide. Pediatr Discov. 2024. PMCID: PMC13612974.

2. Evaluation of a commercial pediatric sepsis early detection alert in the emergency department. JAMA Pediatr. 2026. PMID: 42832216.

3. The burden and contemporary epidemiology of sepsis in children. Lancet Child Adolesc Health. 2024;8(9):683-694.

4. Global, regional, and national sepsis incidence and mortality, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet Glob Health. 2025.

5. Schlapbach LJ, Watson RS, Sorce LR, et al. International consensus criteria for pediatric sepsis and septic shock. JAMA. 2024;331(8):665-674.

6. Sanchez-Pinto LN, Bennett TD, DeWitt PE, et al. Development and validation of the Phoenix criteria for pediatric sepsis and septic shock. JAMA. 2024;331(8):675-686.

7. National estimates of pediatric sepsis in US hospitals using clinical data. Pediatrics. 2026. PMID: 41865411.

8. Etiology, prevalence, and mortality of sepsis among children under five years in Africa: a systematic review and meta-analysis. BMC Infect Dis. 2026;26:13350.

9. The burden of neonatal sepsis and its risk factors in Africa: a systematic review and meta-analysis. BMC Public Health. 2025;25:22076.

10. Early-onset sepsis in preterm neonates of 22-28 weeks' gestation: an international cohort study. J Pediatr. 2026.

11. Multidrug-resistant sepsis in special newborn care units in five district hospitals in India: a prospective cohort study. Lancet Glob Health. 2025;13(5).

12. Beyond early- and late-onset neonatal sepsis definitions. Pediatr Infect Dis J. 2024;43(12).

13. An interpretable machine learning model using routine clinical indicators for early pediatric sepsis risk stratification: a multicenter retrospective cohort study. Front Cell Infect Microbiol. 2026;16:1933201.

14. Development and external validation of age-specific predictive models for sepsis in critically ill neonates and children using multidimensional hematological parameters. Crit Care. 2026;30:6261.

15. Construction and efficacy evaluation of a model for early diagnosis of pediatric sepsis based on LASSO-logistic regression. 2025. PMID: 40933706.

16. Predicting pediatric sepsis and mortality using wearable device data and machine learning in Bangladesh. 2025. PMID: 41252746.

17. A machine learning model for prediction of early-onset neonatal sepsis in low-income and middle-income countries. BMJ Paediatr Open. 2026;10(1):e003561.

18. The Phoenix criteria and other severity scores in pediatric sepsis. JAMA Netw Open. 2026;9(6).

19. External validation of the Phoenix sepsis score in a paediatric intensive care unit in Saudi Arabia. Sci Rep. 2026;16:48807.

20. Aligning prediction models with clinical information needs: infant sepsis case study. JAMIA Open. 2025;8(2):ooaf015.

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved

21. Derivation and validation of predictive models for early pediatric sepsis. JAMA Pediatr. 2025. doi:10.1001/jamapediatrics.2025.3167

22. A scoping review on pediatric sepsis prediction technologies in healthcare. npj Digit Med. 2024;7:342.

23. Clinical prediction models to diagnose neonatal sepsis in low-income and middle-income countries: a scoping review. 2025. PMID: 40204466.

24. Artificial intelligence in pediatric intensive care units: current applications in sepsis management. World J Pediatr. 2026.

25. Microvascular phenotypes in pediatric sepsis identified by machine learning. Crit Care. 2026;30.

26. Association between serum endothelial glycocalyx biomarkers and prognosis in paediatric sepsis. BMC Infect Dis. 2026;26:12688.

27. Biomarkers of glycocalyx injury and endothelial activation in sepsis: systematic review and meta-analysis. J Intensive Care Med. 2023;38(10).

28. Aberrant STAT signaling and T cell dysregulation define a targetable pediatric sepsis endotype. J Clin Invest. 2026;136(14):e202867.

29. Identification and transcriptomic assessment of latent profile pediatric septic shock phenotypes. Crit Care. 2024;28:241.

30. Predictive value of dynamic plasma biomarkers for clinical outcomes in pediatric sepsis. 2025. PMID: 41636191.

31. Serum proteomics in newborns reveals a multi-pathway molecular panel for neonatal sepsis. Sci Rep. 2026;16:78660.

32. Immunomodulation in pediatric sepsis: a narrative review. J Clin Med. 2025;14(9):2983.

33. Molecular mechanisms underlying inflammation in early-onset neonatal sepsis. 2025. PMC12347463.

34. Sepsis-related immunosuppression: a comprehensive review of current research breakthroughs. Crit Care. 2026;30.

35. The Phoenix criteria and other severity scores in identifying pediatric sepsis. PMC13231296.

36. Surviving Sepsis Campaign international guidelines for the management of sepsis and septic shock in children 2026. Intensive Care Med. 2026. doi:10.1007/s00134-026-08360-2

37. Balanced fluid or 0.9% saline in children treated for septic shock (PRoMPT BOLUS). N Engl J Med. 2026. doi:10.1056/NEJMoa2601969

38. Balanced crystalloids versus normal saline for fluid resuscitation in pediatric septic shock: an updated systematic review and meta-analysis. 2026. PMC13719638.

39. Epinephrine vs norepinephrine as initial treatment in children with septic shock. JAMA Netw Open. 2025;8(4).

40. Management and outcome of paediatric septic shock requiring vasoactive therapy: a multicentre prospective observational study. 2026. PMID: 43012742.

41. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.

42. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med. 2019;170(1):51-58.

Copyright © 2026 by the Society of Averroes-EMJ All Rights Reserved