

**UNIVERSITÉ LILLE**

**THESE DE DOCTORAT D'UNIVERSITE**

**Discipline : Epidémiologie pédiatrique**

Présentée par Marie Lampin

**Validation d'un score d'alerte et caractérisation  
des trajectoires de gravité des patients  
hospitalisés dans les unités de surveillance  
continue pédiatriques**

Soutenu le 18 Décembre 2019

Dirigé par le Professeur Alain DUHAMEL

Membres du jury :

<b>Madame le Professeur Nathalie BEDNAREK</b>	Rapporteur
<b>Monsieur le Professeur Bertrand GUIDET</b>	Rapporteur
<b>Monsieur le Professeur Stephane LETEURTRE</b>	Examineur
<b>Monsieur le Professeur Alain DUHAMEL</b>	Examineur, Directeur de thèse

Ce travail a été réalisé au sein de l'équipe d'accueil EA2694, dirigé par Monsieur le Professeur Alain DUHAMEL.

## Laboratoire d'accueil : EA 2694

Pr Alain DHAMEL

EA2694 : Santé Publique épidémiologie et qualité des soins

CERIM, Faculté de Médecine

Tel: +33 (0) 3-20-62-69-69

## RESUME

### **Validation d'un score d'alerte et caractérisation des trajectoires de gravité des patients hospitalisés dans les unités de surveillance continue pédiatriques**

**Contexte :** Les unités de surveillance continue (USC) pédiatriques constituent des structures de soins intermédiaires, entre les services « classiques » de pédiatrie (dits « soins courants ») et les services de réanimation, pour les enfants nécessitant une surveillance continue sans mise en œuvre de méthode de suppléance. Ces patients sont à haut risque de dégradation et la validation d'un score d'alerte précoce (SAP) dans cette population serait intéressante.

**Objectifs :** 1- Valider un score d'alerte précoce en USC pédiatriques.  
2- Classifier les malades de ces USC pédiatriques selon leurs caractéristiques générales, leurs affections et les trajectoires de gravité de leur maladie.

**Méthodes :** Etude observationnelle prospective multicentrique régionale dans sept USC pédiatriques françaises incluant tous les enfants consécutifs admis de septembre 2012 à janvier 2014. Validation de trois SAP, recueillis toutes les 8 heures (le Pediatric Advanced Warning Score (PAWS), le Pediatric Early Warning Score (PEWS) et le Bedside Pediatric Early Warning System (Bedside PEWS)) dans la population d'USC en utilisant un modèle linéaire mixte généralisé pour mesures répétées. La cohorte était divisée en échantillon de dérivation (70%) et en échantillon de validation (30%). La discrimination de ces SAP pour la prédiction de l'appel du médecin par l'infirmière en cas de détérioration clinique était mesurée par l'aire sous la courbe ROC. Un modèle mixte linéaire à classes latentes était utilisé pour identifier différentes trajectoires de gravité de ces patients d'USC.

**Résultats :** Un total de 2868 enfants a été inclus pour 14708 observations utilisables pour calculer a posteriori les SAP. La discrimination des trois SAP pour prédire l'appel du médecin était bonne (entre 0,87 et 0,91) sur la cohorte de dérivation et modérée (entre 0,71 et 0,76) sur la cohorte de validation.

La défaillance principale à l'origine de l'admission en USC était respiratoire (44%) et l'étiologie infectieuse était la plus fréquente (52%). Les deux diagnostics les plus fréquents étaient l'asthme et la bronchiolite. Dix diagnostics représentaient 58% des patients d'USC. La médiane de durée de séjour était de 1 jour [1-3]. L'analyse en classe latente mettait en évidence différentes trajectoires de gravité : le profil « stable » (60,4%), le profil « amélioration rapide » (6,5%) et le profil « amélioration lente » (33,1%).

**Conclusion :** Les SAP peuvent être utilisés dans les USC pour détecter une détérioration clinique et prédire la nécessité d'une intervention médicale. Trois trajectoires très différentes de gravité ont été identifiées avec une majorité de patients appartenant au profil « stable ».

Mots clefs: enfant, unités de surveillance continue, score d'alerte précoce, détérioration clinique, performance, gravité, modèle linéaire mixte généralisé, classes latentes

## ABSTRACT

### **Validation of an early warning score and characterization of trajectory-focused patient profiles in pediatric intermediate care units**

Background: Pediatric Intermediate care units (PImCU) are Intermediate care units (ImCU) or high dependency care units (HDC), between regular wards and intensive care units (ICUs), for children requiring continuous monitoring without active life-supporting treatment. These patients are at high risk of deterioration and the validation of an early warning score (EWS) in this population would be interesting.

Objectives: 1- To validate early warning scores in PImCU

2- To classify patients according to their general characteristics, their diagnoses and the severity trajectories of illness.

Methods: Regional multicenter prospective observational study in seven French PImCU including all consecutive children admitted from September 2012 to January 2014. Validation of EWS in PImCU using a general linear mixed model for repeated measures. The cohort was divided into derivation (70%) and validation (30%) cohorts. The discrimination to predict physician call by nurse was estimated by the area under the receiver-operating curve. A latent class linear mixed model was used to identify different trajectories of severity of illness of PImCU patients.

Results: A total of 2868 children were included for 14708 observations to compute a posteriori the EWS. The discrimination of the three EWS for predicting calls to physicians by nurses was good (range: 0.87–0.91) for the derivation cohort and moderate (range: 0.71–0.76) for the validation cohort. The primary failure for admission to PImCU was respiratory (44%) and infectious etiology was the most common (52%). The two most common diagnoses are asthma and bronchiolitis. Ten diagnoses account for 58% of PImCU patients. The median length of stay was 1 day [1-3]. The latent class analysis identified different trajectories of severity of illness: profile "stable" (60.4%), profile "rapid improvement" (6.5%) and profile "slow improvement" (33.1%).

Conclusion: SAP can be used in PImCU to detect clinical deterioration and predict the need for medical intervention. Three very different trajectories of severity were identified with a majority of "stable" profile.

**Keywords:** child, intermediate care units, early warning score, clinical deterioration, performance, severity of illness, latent classes

## REMERCIEMENTS





## VALORISATION SCIENTIFIQUE

### **Publications**

#### Acceptées:

*Les scores de gravité généraux et de dysfonctions d'organes en réanimation pédiatrique : quoi de neuf en 2016 ?*

S. Leteurtre, M.-E. Lampin, B. Grandbastien, M. Recher, A. Duhamel.  
Méd. Intensive Réa (2016) 25:604-618.

*Use of paediatric early warning scores in intermediate care units*

Marie Emilie Lampin, Alain Duhamel, Hélène Behal, Morgan Recher, Francis Leclerc, Stéphane Leteurtre. Arch Dis Child 2019;0:1–6.

#### A soumettre:

*Trajectory-Focused Patient Profiles and Patient Characteristics in Pediatric Intermediate Care Unit.*

Marie Emilie Lampin, Alain Duhamel, Hélène Behal, Francis Leclerc, Stéphane Leteurtre

### **Communications**

#### Orales:

*Recherche d'un indicateur de gravité et d'activité simplifié pour les unités de surveillance continue pédiatriques : étude prospective multicentrique régionale*

Conférence de Recherche en Soins : « de la recherche en soins au changement des pratiques ».

Marie Emilie Lampin, Isabelle Grit, Stéphane Leteurtre

Organisé par le CHU de Lille, 4 décembre 2018, Institut Gernez Rieux

#### Affichées:

Validation of three Pediatric Early Warning Scores in pediatric intermediate care unit.

Lampin ME, Duhamel A, Béhal H, Dhaoui T, Goddefroy V, Pouessel G, Devouge E, Evrard D, Delepouille F, Racoussot S, Grandbastien B, Leclerc F, Leteurtre S

8th World Congress on Pediatric Intensive and Critical Care  
Toronto, Canada, 04-08/06/2016

Caractéristiques des patients hospitalisés en unité de surveillance continue pédiatrique

Lampin ME, Duhamel A, Béhal H, Dhaoui T, Goddefroy V, Pouessel G, Devouge E, Evrard D, Delepouille F, Racoussot S, Grandbastien B, Leclerc F, Leteurtre S

Congrès Réanimation, organisé par la société de réanimation de langue française, Paris, 11-15/01/2017.

Validation of three Pediatric Early Warning Scores in seven French pediatric intermediate care units

Lampin ME, Duhamel A, Béhal H, Dhaoui T, Goddefroy V, Pouessel G, Devouge E, Evrard D, Delepouille F, Racoussot S, Grandbastien B, Leclerc F, Leteurtre S

Congrès Réanimation, organisé par la société de réanimation de langue française, Paris, 11-15/01/2017.

## ABBREVIATIONS

USC : unités de surveillance continue

DHOS : direction de l'hospitalisation et de l'organisation des soins

PIM : Paediatric Index of Mortality

PRISM : Pediatric RISK of Mortality

DO : dysfonctions d'organes

PELOD : PEdiatric Logistic Organ Dysfunction

AUC : Area Under the Receiver Operating Curve

RSM : ratio standardisé de mortalité

SMR : Standardized Mortality Ratio

SAP: scores d'alerte précoce

SRR: système de réponse rapide

SAPS II : Simplified Acute Physiology Score II

SAPS 3 : Simplified Acute Physiology Score 3

ImCUSS : Intermediate Care Unit Severity Score

PAWS : Pediatric Advanced Warning Score

PEWS : Pediatric Early Warning Score

Bedside PEWS: Bedside Pediatric Early Warning System

CHG : centres hospitaliers généraux

CHU : centre hospitalier universitaire

EPILS-EPALS : European Paediatric Immediate and advanced Life Support

PHRC : Programme Hospitalier de Recherche Clinique

CTD : Centre de Traitement des Données OR : odds ratio

CNIL : Commission Nationale d'Informatique et Liberté

CCTIRS: comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé

SRLF: société de réanimation de langue française

CDR PEWS: Central Denmark Region Pediatric Early Warning System

CCAM: classification commune des actes médicaux

DIM : Département d'Informatique Médical

LCLMM : modèle mixte linéaire à classes latentes

AIC : Akaike information criterion

BIC : Bayes Information Criterion

ANOVA : analyse de la variance

CIM10 : Classification international des maladies 10<sup>ème</sup> révision

# Table des matières

<b>Introduction.....</b>	<b>14</b>
1- Les unités de surveillance continue pédiatrique.....	14
2- Evaluation de l'état de gravité des patients .....	15
2.1- Généralités .....	15
2.2- Les scores de gravité généraux et de dysfonctions d'organes utilisés en réanimation pédiatrique.....	16
2.3- Les scores d'alerte précoce en pédiatrie .....	17
2.4- Quels scores de gravité utilisés pour les unités de surveillance continue.....	19
<b>Objectifs du travail.....</b>	<b>21</b>
<b>Validation d'un score d'alerte précoce au sein des unités de surveillance continue pédiatriques.....</b>	<b>22</b>
1- Méthodologie.....	22
1.1- Schéma de l'étude .....	22
1.2- Adhésion au projet des centres participants.....	22
1.3- Population et durée de l'étude.....	23
1.4- Critère de jugement pour la validation du score .....	23
1.5- Etude de reproductibilité (validité inter-observateur).....	24
1.6- Origine et nature des données recueillies.....	24
1.7- Calcul du nombre de sujet nécessaire .....	27
1.8- Analyses statistiques .....	27
1.9- Considérations éthiques .....	28
2- Résultats.....	30
3- Discussion.....	36
3.1- Choix du critère de jugement.....	36
3.2- Choix d'un modèle linéaire mixte généralisé pour mesures répétées	
3.3- Proposition d'utiliser les scores d'alerte précoce dans une équation prédictive d'appel du médecin .....	38
3.4- Evolutions de la performance des scores au cours de l'hospitalisation.....	39
3.5- Reproductibilité des scores d'alerte précoce .....	40
3.6- Gestion des données manquantes .....	40
3.7- L'effet centre .....	41
3.8- Représentativité des patients d'une USC non attachée à une réanimation pédiatrique.....	42
3.9- Quelle conclusion de notre article et quel score d'alerte précoce choisir en pratique?.....	42

1-	Méthodologie .....	44
1.1-	Schéma de l'étude et population .....	44
1.2-	Données utilisées .....	44
1.3-	Critère de jugement.....	46
1.4-	Choix du SAP .....	46
1.5-	Analyses statistiques .....	47
2-	Résultats.....	48
3-	Discussion.....	59
3.1-	Caractéristiques générales, défaillances aiguës et étiologies.....	59
3.2-	Les patients transférés en réanimation.....	60
3.3-	L'analyse en classe latente.....	61
3.4-	Limites .....	62
	<b>Conclusion .....</b>	<b>63</b>
	<b>Références.....</b>	<b>64</b>
	<b>Annexes .....</b>	<b>70</b>
	Annexe 1: The Paediatric advanced warning score (PAWS), Egdell et al.....	70
	Annexe 2: The Paediatric Early Warning System (PEWS), Tucker et al .....	71
	Annexe 3: The Bedside paediatric early warning system score (Bedside PEWS), Parshuram et al.....	72
	Annexe 4: Use of Paediatric Early Warning Scores in intermediate care units. 73	
	Annexe 5: Fiche de recueil des caractéristiques générales.....	101
	Annexe 6 : Feuille de surveillance standardisée.....	102
	Annexe 7 : Gestion des données manquantes.....	103
	Annexe 8: Note d'information pour les parents de patients .....	104
	Annexe 9: Note d'information pour les enfants .....	106
	Annexe 10: formulaire d'opposition à remplir par le médecin ou le soignant 108	
	Annexe 11: Description du modèle mixte linéaire généralisé utilisé pour la validation des scores d'alerte précoce .....	109
	Annexe 12: Descriptif des données manquantes .....	110
	Annexe 13 : Effet centre .....	111
	Annexe 14 : Données DIM .....	112

# Introduction

## 1- Les unités de surveillance continue pédiatrique

Les unités de surveillance continue (USC) pédiatriques constituent des structures de soins intermédiaires entre les services « classiques » de pédiatrie (dits « soins courants ») et les services de réanimation<sup>1-3</sup>. L'organisation de ces USC pédiatriques est régie par les décrets 2006-72 et 2006-48 de janvier 2006 et la circulaire de la direction de l'hospitalisation et de l'organisation des soins<sup>4</sup>. Ce décret stipule que les USC pédiatriques sont destinées aux enfants «qui nécessitent une surveillance rapprochée ou continue en raison d'un risque de défaillance d'un ou de plusieurs organes ne nécessitant pas la mise en œuvre de méthodes de suppléance ». Ces USC permettent ainsi de diminuer le nombre de transferts vers le service de réanimation pédiatrique.

Les critères d'admission et de sortie de ces USC restaient à définir plus précisément pour veiller à ce que ces structures répondent bien à l'objectif qui leur est assigné<sup>3</sup>. L'Académie américaine de pédiatrie a proposé ainsi des critères assez généraux médicaux ou chirurgicaux, selon chaque spécialité d'organe, avec les conditions d'admission et de sortie de ces USC pédiatriques<sup>2</sup>. Le fonctionnement de ces USC est abordé dans le décret<sup>3,4</sup>: l'article R. 6123-38-7 stipule que les USC ont pour mission de prendre en charge les nourrissons, enfants et adolescents qui nécessitent une surveillance rapprochée ou continue en raison d'un risque de défaillance d'un ou de plusieurs organes ne nécessitant pas la mise en œuvre de méthode de suppléance. Elles peuvent être médicale et chirurgicale. Lorsque l'établissement dispose d'une réanimation pédiatrique, l'USC est associée à la réanimation et est individualisée au sein de celle-ci ou à proximité immédiate. Un établissement de santé ne disposant pas de réanimation pédiatrique peut disposer d'une USC s'il a conclu une convention de transferts des enfants avec un établissement de santé qui dispose d'une réanimation pédiatrique. Sa capacité et son organisation peuvent varier en fonction du type d'établissement mais répondent à certains critères minimaux : l'équipement de l'USC pédiatrique doit permettre, au niveau de chaque lit, un monitoring continu non invasif des principales constantes vitales. L'équipe médicale intervenant dans ces unités est composée de pédiatres et d'anesthésistes réanimateurs. La permanence médicale peut être mutualisée avec le service de pédiatrie, les urgences pédiatriques ou la néonatalogie ou la réanimation pédiatrique lorsqu'il y en a une. Elle est constituée au minimum d'une astreinte opérationnelle. Concernant le

personnel paramédical, il est recommandé que l'USC bénéficie de la présence d'un infirmier pour quatre patients.

L'article D. 6124-120 précise que lorsque l'établissement dispose de réanimation pédiatrique, la responsabilité médicale est commune pour l'USC et la réanimation pédiatrique. Lorsque l'établissement ne dispose pas de réanimation pédiatrique, la responsabilité médicale de l'USC peut être commune avec celle du service de pédiatrie, de néonatalogie ou des urgences pédiatriques. Ces USC pédiatriques se sont largement développés ces deux dernières décennies mais les critères d'admission et de sortie n'ont pas été plus précisément définis. Il existe peu de données dans la littérature sur l'activité de ces USC pédiatriques et sur les caractéristiques des patients qui y sont hospitalisés.

Actuellement, les ressources allouées à ces USC sont déterminées à partir de 3 listes de diagnostics et d'actes associés (2 listes adultes et une liste pédiatrique) autorisant la facturation d'un supplément de surveillance continue (Arrêté du 22 février 2019). Ces listes pour essayer d'être exhaustives contiennent une multitude de diagnostics et d'actes n'étant pas toujours synonyme de besoin de surveillance continue. Lors d'une réunion à la DHOS le 23 novembre 2009, rassemblant l'ensemble des Sociétés professionnelles, il a été discuté l'utilité d'avoir un indicateur qui devrait cerner la quantité de soins et/ou de surveillance prodigués pour ces USC. Afin de mieux cerner la quantité de soins ou de surveillance nécessaire, cet indicateur devrait prendre en compte le niveau de gravité de l'état de l'enfant.

## **2- Evaluation de l'état de gravité des patients**

### **2.1- Généralités**

L'évolution d'un patient et de sa maladie peut être favorable ou défavorable et dans les 2 cas, cette évolution peut être lente ou rapide. L'évaluation de la gravité d'un patient fait partie intégrante de l'activité quotidienne du médecin. La gravité de l'affection est variable au cours du temps et dépendra des antécédents du patient, de l'état actuel de celui-ci, de sa réserve physiologique, des traitements entrepris. Lorsqu'un patient est admis, son pronostic dépend de facteurs présents à l'admission mais aussi des événements qui vont survenir pendant son hospitalisation.

Chez l'enfant, l'évaluation de l'état de gravité peut être difficile du fait des conditions d'examen médical (absence de compliance, absence de participation, peur du soignant, absence de verbalisation des symptômes) et de la variabilité des normes

physiologiques selon l'âge. De plus, chez l'enfant, l'état clinique peut se détériorer ou s'améliorer plus rapidement que chez l'adulte. Il existe plusieurs systèmes d'évaluation de la gravité en pédiatrie, adaptés au risque encouru par l'enfant.

## **2.2- Les scores de gravité généraux et de dysfonctions d'organes utilisés en réanimation pédiatrique**

### 2.2.1- Descriptif de ces scores de gravité généraux et de dysfonctions d'organes pédiatriques

Dans les services de réanimation, le risque encouru est le décès. Ainsi, les scores de gravité généraux développés pour les enfants hospitalisés en réanimation, tels que les scores Paediatric Index of Mortality (PIM)<sup>5</sup> et Pediatric RISK of Mortality (PRISM)<sup>6</sup>, ont été construits en prenant la mortalité comme critère de jugement. Ces scores sont des scores pronostiques, prédictifs de mortalité. Ils se calculent une fois au cours du séjour, durant la première heure après l'admission pour le PIM et dans les 12 à 24 premières heures pour le PRISM. Parallèlement à ces scores de gravité généraux, il existe des scores de dysfonctions d'organes (DO) comme le PEdiatric Logistic Organ Dysfunction (PELOD) en pédiatrie<sup>7</sup>. Ces scores ont pour but de décrire les DO à l'admission mais aussi tout au long du séjour. Le critère de jugement pour le développement et la validation des scores de DO (PELOD et PELOD-2 en pédiatrie) est également la mortalité. Ils peuvent être recueillis quotidiennement mais sont alors très consommateurs de ressources médicales. Des jours dits « ciblés » pour le recueil du score PELOD les plus liés à la mortalité durant le séjour en réanimation ont été déterminés<sup>7</sup>. Une mise au point, à laquelle j'ai participé, a été publiée en 2016 sur les caractéristiques des scores de gravités généraux et les scores de DO en pédiatrie<sup>8</sup>

### 2.2.2- Méthodologie pour le développement et la validation de ces scores

Cette mise au point détaillait les méthodologies utilisées pour le développement et la validation des scores de gravité et de DO<sup>8</sup>.

Pour le développement de ces scores, le critère de jugement utilisé est la mortalité : c'est un critère binaire (oui/non) de type « gold standard ». Les qualités requises de ces scores de gravité sont multiples. Les variables du score doivent être pertinentes, usuelles, objectives, de recueil facile. Le score doit avoir une bonne reproductibilité intra et inter-observateur. Une validation interne et externe est nécessaire. La validation interne consiste à déterminer les qualités du score lorsqu'il est appliqué à l'échantillon qui a

permis son développement. La validité externe consiste à déterminer ses qualités lorsqu'il est appliqué à un autre échantillon. Il doit être « acceptable » pour le patient et facile d'utilisation pour le médecin<sup>9</sup>. Les outils statistiques utilisés pour développer et évaluer la performance de ces scores sont le pouvoir discriminant et la calibration<sup>10,11</sup>. La discrimination d'un score de gravité est sa capacité à séparer les sujets qui vont vivre de ceux qui vont décéder. La discrimination est évaluée par le calcul de l'aire sous la courbe ROC (Area Under the Receiver Operating Curve –AUC- des Anglo-saxons). L'AUC est obtenue en faisant varier successivement les seuils du score et en calculant pour chacun les sensibilités et spécificités. Une AUC égale à 0,50 n'est pas plus discriminante que le hasard, une AUC entre 0,70 et 0,79 est correcte, une AUC entre 0,80 et 0,89 est bonne et une AUC > à 0,90 est excellente<sup>10,12,13</sup>. La calibration du score est l'adéquation entre la mortalité prédite par le score et la mortalité observée. Elle est évaluée par le test de Hosmer-Lemeshow qui effectue une comparaison des nombres de décès observés et de décès prédits (test du Chi-2) par groupe de niveaux de gravité (ou test de goodness-of-fit des anglosaxons)<sup>14</sup>. Une valeur de p supérieure à 0,05 témoigne de cette adéquation entre la mortalité prédite et la mortalité observée. Pour étudier si le score est capable de prédire le nombre de décès dans un service, on peut calculer le ratio standardisé de mortalité (RSM) (ou Standardized Mortality Ratio –SMR), défini comme le rapport entre le nombre de décès observés divisé par le nombre de décès prédits. Lorsque le SMR est inférieur à 1, le nombre de décès observés est inférieur au nombre de décès prédits et s'il est supérieur à 1, le nombre de décès observés est supérieur au nombre de décès prédits. Si l'intervalle de confiance du SMR inclut 1, la différence entre le nombre de décès observés et le nombre de décès prédits n'est pas significative<sup>8</sup>.

## **2.3- Les scores d'alerte précoce en pédiatrie**

### **2.3.1- Descriptif de ces scores d'alerte précoce**

Pour les patients hospitalisés dans les services « classiques » de pédiatrie et également aux urgences, de nouveaux outils d'évaluation de la gravité ont été développés ces dernières années<sup>15</sup>. Bien que le nombre d'arrêt cardio-respiratoire survenant chez des enfants hospitalisés est relativement faible, évalué à 0,19 à 2,45 pour 1000 admissions, la mortalité et la morbidité associée reste élevée<sup>16</sup>. Il a été suggéré qu'un système d'alerte précoce (tel un score) utilisable au lit du malade et permettant de détecter la survenue d'une détérioration clinique serait utile et permettrait d'éviter ces arrêts cardio-respiratoires en permettant une intervention précoce adaptée. Le but de ces scores d'alerte

précoces (SAP) étant de détecter la survenue d'une détérioration, les critères de jugement utilisés pour leur développement et leur validation sont nombreux et variés. Il n'existe pas de consensus sur le gold standard établissant la détérioration clinique<sup>17</sup>. La construction de ces SAP est donc basée sur différents critères de jugement : décès<sup>18</sup>, arrêt cardiaque ou code bleu<sup>19-21</sup>, transfert non planifié vers la réanimation pédiatrique<sup>21-27</sup>, appel à une assistance médicale urgente ou activation du système de réponse rapide (SRR)<sup>21,24,28-31</sup> et durée du séjour à l'hôpital<sup>28</sup>. Ces différents SAP sont constitués de paramètres physiologiques de surveillance tels que la fréquence cardiaque, la pression artérielle, la fréquence respiratoire, l'effort respiratoire. Ils ont une diversité considérable dans le nombre, le type des paramètres contrôlés, ainsi que dans les seuils retenus pour les variables numériques et les différentes classes d'âge utilisées pour certains items dépendants de l'âge. Tous incluent la surveillance des signes vitaux et certains y ajoutent des paramètres liés au diagnostic ou à une intervention.

### 2.3.2- Méthodologie pour le développement et la validation des scores d'alerte précoce

Chez l'adulte, les SAP pour détecter la détérioration sont également nombreux et sont largement utilisés. Des données récentes suggèrent que les SAP ne répondent pas au problème pour lequel ils ont été développés et que la reconnaissance de la détérioration des patients dans les hôpitaux reste un problème majeur<sup>33</sup>. Une revue récente s'est intéressée aux SAP en mettant l'accent sur la méthodologie utilisée pour l'élaboration et la validation de ces SAP et sur leur généralisabilité, suggérant que les méthodes utilisées pour leur développement et leur validation manquaient de robustesse et pourraient être responsable du manque de fiabilité et de la mauvaise prédiction du risque de ces SAP<sup>32</sup>. La méthodologie recommandée pour un score prédictif comme le SAP, dans son développement et sa validation devrait aussi inclure l'étude de la discrimination et la calibration<sup>33</sup>, comme nous l'avons vu précédemment avec les scores de gravité généraux et de DO pédiatriques, développés pour les réanimations.

Chez les enfants, de nombreux SAP ont également été développés et testés chez les patients hospitalisés en services classiques ou aux urgences. Trois revues systématiques des SAP ont été publiées<sup>15,17,34</sup>. Pour le développement ou la validation des SAP en pédiatrie, on peut faire le même constat sur le manque de robustesse et de fiabilité. Aucun n'a été développé ou validé en utilisant la méthodologie recommandée (étude de la discrimination et de la calibration).

### 2.4- Quels scores de gravité utilisés pour les unités de surveillance continue

Chez l'adulte, le taux de mortalité dans les USC a été moins rapporté que la mortalité à l'hôpital après le séjour en USC. Certaines études ont observé un taux de mortalité à la sortie de l'USC de 2,7%, 5,1% et 5,9% respectivement<sup>35-37</sup>, tandis que la mortalité à la sortie de l'hôpital était plus élevée de 8,3%, 8,9%, 19,7% et 20,6%<sup>35,37-39</sup>. Certains scores de gravités généraux développés en réanimations adultes, utilisant comme critère de jugement la mortalité comme le Simplified Acute Physiology Score II ou 3 (SAPS II ou 3)<sup>40,41</sup>, ont été testés en USC adultes<sup>35,42</sup>. Dans ces deux études, le pouvoir discriminant du score était évalué par l'AUC et la calibration du score par le test statistique de Goodness of Fit de Hosmer-Lemeshow. Dans la première étude, sur une cohorte prospective de 433 patients consécutifs inclus, le pouvoir discriminant du score SAPS II en USC était bon (AUC de  $0,85 \pm 0,04$ ) et la calibration était bonne ( $p > 0,05$ )<sup>35</sup>. L'étude était réalisée dans un hôpital universitaire dont l'USC était à proximité d'une réanimation. Dans la seconde étude incluant 607 patients d'USC, attachée également à

une unité de réanimation, le SAPS II et 3 étaient testés. Les deux scores avaient une performance similaire (AUC correcte de 0,76 et 0,75 respectivement et une calibration bonne avec  $p > 0,05$  pour les deux scores)<sup>42</sup>. Alegre et *al.*, ont proposé de développer un score de sévérité spécifiquement pour les USC adultes<sup>39</sup>. A partir d'une cohorte rétrospective de 743 patients admis en USC, attachée à un service de réanimation, une régression logistique a été réalisée comprenant des données démographiques, les antécédents, les raisons de l'admission, paramètres physiologiques à l'admission et durant les premières 24 heures du séjour, paramètres biologiques afin d'identifier les variables prédictives de mortalité sur une cohorte de développement (comprenant le 2/3 de la cohorte,  $n=554$ ). Le score obtenu Intermediate Care Unit Severity Score (ImCUSS) comprenait 9 variables. L'équation de calcul de la probabilité de décès était fournie. Le score était ensuite validé sur une cohorte de validation (le 1/3 restant,  $n=189$ ). La discrimination (AUC=0,802) et la calibration ( $p > 0,05$ ) de l'ImCUSS étaient bonnes. Hager et *al.*, ont évalué ce nouveau score développé spécifiquement pour les USC dans une étude rétrospective incluant 682 patients admis en USC. La discrimination du score était correcte (AUC = 0,72)<sup>38</sup>.

Pour les patients hospitalisés dans les USC pédiatriques, aucun des scores de gravité disponibles pour la réanimation pédiatrique (PRISM ou le PIM dans leur différente version) n'a pour le moment été testé (la mortalité en USC pédiatrique ou à la sortie de l'hôpital étant plus faible que dans la population adulte). Aucun score n'a été développé spécifiquement pour les USC pédiatriques. Un tel score devrait utiliser comme critère de jugement la nécessité d'une admission en réanimation ou la nécessité d'une évaluation médicale, plutôt que la mortalité, événement trop rare en pédiatrie. Ces critères de jugement, utilisés pour le développement ou la validation des SAP sur des patients de soins courants ou aux urgences, pourraient être, malgré tout, adaptés pour détecter la détérioration des patients d'USC pédiatriques. La description des SAP pédiatriques est effectuée dans le premier axe du travail (après le paragraphe objectif). Les 3 principaux scores que nous avons utilisés sont le Pediatric Advanced Warning Score (PAWS)<sup>43</sup>, le Pediatric Early Warning Score (PEWS)<sup>22</sup> et le Bedside Pediatric Early Warning System (Bedside PEWS)<sup>23</sup> (Annexes 1, 2 et 3).

## Objectifs du travail

L'objectif du travail se décline selon deux axes :

### **Axe 1 : Valider un score d'alerte précoce en unités de surveillance continue pédiatriques**

Nous avons fait l'hypothèse que les SAP seraient pertinents pour refléter la gravité des patients hospitalisés en USC pédiatriques. Nous avons choisi d'évaluer les performances de trois SAP existants, le PAWS<sup>43</sup>, le PEWS<sup>22</sup> et le Bedside PEWS<sup>23</sup> au sein des USC pour détecter la détérioration clinique. Les trois SAP évalués sont présentés Annexe 1, 2 et 3.

### **Axe 2 : Caractérisation des trajectoires de gravité des patients d'unités de surveillance continue pédiatriques**

Le deuxième axe était d'identifier différents profils de patients en fonction des trajectoires de gravité, et d'identifier les caractéristiques et les diagnostics principaux associés à ces différents profils de gravité.

# Validation d'un score d'alerte précoce au sein des unités de surveillance continue pédiatriques

Ce premier objectif a abouti à la publication d'un article, accepté par la revue Archives of Disease in Childhood (Annexe 4).

Il existe une multitude de SAP en pédiatrie. Nous avons choisi d'évaluer trois de ces SAP, le PAWS<sup>43</sup>, le PEWS<sup>22</sup> et le Bedside PEWS<sup>23</sup> au sein des USC pour détecter la détérioration clinique (Annexe 1,2 et 3).

Les trois SAP sélectionnés sont faciles à utiliser et réalisables au lit du patient. Deux d'entre eux incluent sept paramètres de surveillance. Le PAWS<sup>43</sup> a été mis au point dans les services d'urgences et le Bedside PEWS<sup>23</sup> a été mis au point dans les services de pédiatrie hospitalière. Le PAWS comprend une évaluation neurologique contrairement au Bedside PEWS. Le Bedside PEWS a, par contre, l'avantage de prendre en compte la saturation en oxygène et les apports en oxygène, contrairement au PAWS, qui ne considère que la saturation en oxygène. Nous avons sélectionné un troisième SAP, le PEWS<sup>22</sup>, testé par plusieurs équipes en version modifiée<sup>19,22,28,44,45</sup>. Nous avons choisi ce troisième SAP car il est présenté en terme de défaillances d'organes (respiratoire, circulatoire et neurologique) comme présenté dans les recommandations de l'European Resuscitation Council.

## 1- Méthodologie

### 1.1- Schéma de l'étude

Ce travail s'inscrit dans le cadre d'une étude observationnelle prospective multicentrique régionale française incluant les USC de sept centres hospitaliers généraux (CHG) du Nord-Pas de Calais (Arras, Boulogne, Cambrai, Douai, Dunkerque, Lens, Roubaix) et le service de réanimation pédiatrique du centre hospitalier universitaire (CHU) de Lille, unique structure de ce type pour le Nord-Pas de Calais et service de recours de ces 7 USC.

### 1.2- Adhésion au projet des centres participants

L'ouverture des premières USC dans la région Nord-Pas de Calais a suivi les décrets 2006-72 et 2006-48 du 24 janvier 2006 et la circulaire DHOS 0/2006/396 du 8 septembre 2006. L'article 1.6 stipulait que « Afin de mettre en place cette filière de soins les professionnels concernés s'organisent au sein d'un comité. [...] Ce comité peut également

organiser la formation et le maintien des compétences des personnels médical et paramédical travaillant dans les différents services concernés. Il peut être à l'initiative de recherches à visée épidémiologique ou thérapeutique ». Les premières réunions d'organisation de la filière de soins régionale entre les USC et le service de réanimation pédiatrique du CHU de Lille ont eu lieu en 2007. Deux journées de formation (European Paediatric Immediate et Advanced Life\_Support –EPILS et EPALS-) communes aux médecins et paramédicaux ont été effectuées depuis 2010 dans le but d'harmoniser les pratiques de soins. Cette étude multicentrique régionale a également obtenu le soutien du Programme Hospitalier de Recherche Clinique (PHRC), en 2011.

### **1.3- Population et durée de l'étude**

#### 1.3.1- Critères d'inclusion

Tous les enfants admis dans les USC de ces sept CHG ont été inclus dans l'étude.

#### 1.3.2- Critères de non-inclusion

Les enfants dont les parents ou représentant légaux refusaient l'inclusion de leur enfant, étaient exclus de l'étude.

#### 1.3.3- Durée des inclusions

Les inclusions des enfants ont eu lieu entre le 7 septembre 2012 et le 7 janvier 2014. Du fait d'une préparation et d'une formation des équipes nécessaires dans chaque centre, le premier jour d'inclusion pour les sept centres participants a été différent s'étalant du 07 septembre 2012 au 07 janvier 2013. La date de fin d'inclusion était la même pour tous les centres. Tous les centres ont inclus pendant un an minimum.

### **1.4- Critère de jugement pour la validation du score**

Pour répondre aux objectifs de cette étude, un critère de jugement binaire pour valider le score a été retenu : l'appel d'un médecin par l'infirmière (oui/non), lorsqu'elle détecte une détérioration clinique

### **1.5- Etude de reproductibilité (validité inter-observateur)**

Afin d'étudier la reproductibilité interindividuelle des scores d'alerte précoce, à un jour ouvré de la semaine établie au début de l'étude et fixé par les investigateurs de chaque centre (tous les lundi à 10h00 par exemple), le médecin investigateur du centre ou son représentant remplissait un score au même moment que l'infirmière, mais en aveugle l'un de l'autre.

### **1.6- Origine et nature des données recueillies**

Le recueil prenait en compte des données descriptives des patients et des données relatives à l'évaluation de la gravité des patients.

#### **1.6.1- Anonymisation des données**

L'anonymat des patients était conservé grâce à un code d'anonymisation attribué à chaque patient. Une liste établissant la correspondance entre le code d'anonymisation et l'identité de l'enfant participant était renseignée tout au long de l'étude dans chacun des centres. La liste était conservée dans le centre sous la responsabilité du médecin investigateur de chaque centre et était mise à la disposition de l'attaché de recherche clinique.

#### **1.6.2- Données descriptives : caractéristiques générales**

Les données descriptives recueillies comprenaient la date et l'heure d'admission, la date de naissance, le sexe, la provenance de l'enfant (domicile, service d'urgence, service de pédiatrie, autres), les facteurs de morbidité (affection chronique), la nature médicale ou chirurgicale de l'affection, la principale défaillance viscérale entraînant l'admission, la date et l'heure de sortie, la destination de l'enfant, le devenir à la sortie de l'USC (vivant/décédé). La fiche de recueil des caractéristiques générales est présentée en Annexe 5.

#### **1.6.3- Données relatives à l'évaluation de la gravité**

La surveillance habituelle des patients admis en USC comprenait le recueil régulier de différents paramètres cliniques permettant d'apprécier le niveau de gravité de l'état de l'enfant. Ces données étaient recueillies habituellement par les équipes de soins sur des

feuilles de surveillance journalière. Les feuilles de surveillance journalière utilisées par les centres avant l'étude n'étaient pas identiques. Le groupe des investigateurs des différents centres s'est réuni plusieurs fois et a travaillé ensemble pour établir une feuille de surveillance journalière commune (présentée en Annexe 6). Aucune variable supplémentaire n'a été ajoutée pour l'étude, seule la disposition des variables dans la feuille a été harmonisée. Ces variables étaient habituellement recueillies de façon pluriquotidienne dans les USC (toutes les 2 à 4 heures). L'objectif de standardiser cette feuille de surveillance journalière était d'éviter une seconde feuille de recueil qui aurait considérablement alourdi le travail de collecte et de saisie des données. En fin de séjour, une photocopie de ces feuilles de surveillance anonymisée était faite et ces photocopies étaient envoyées au service de réanimation du CHU de Lille (centre coordinateur) et étaient utilisées directement pour la saisie.

Les paramètres cliniques recueillis sur cette feuille de surveillance journalière ont permis le calcul des scores d'alerte précoce à posteriori lors de l'analyse des données. Ainsi les soignants ne disposaient pas de valeur de scores pour prendre une décision.

Ces paramètres cliniques de surveillance collectés étaient au nombre de neuf :

- fréquence respiratoire,
- travail respiratoire,
- saturation en oxygène,
- besoins en oxygène
- température,
- temps de recoloration cutanée,
- fréquence cardiaque,
- pression artérielle systolique/diastolique (moyenne),
- échelle simplifiée de conscience.

Ces neuf paramètres cliniques étaient relevés, à l'admission, puis au minimum toutes les 8 heures, (soit 1 fois par poste d'infirmière). Lorsque l'infirmière était inquiète de l'état clinique de l'enfant et qu'elle jugeait nécessaire une intervention médicale auprès de l'enfant, ces paramètres étaient également recueillis. Il était alors précisé en plus de l'appel du médecin par l'infirmière, la réalisation ou non d'un examen complémentaire demandé par le médecin et/ou la prescription d'un nouveau traitement. Le jour et l'heure d'appel du médecin auprès de l'enfant par l'infirmière était notifié. Ces paramètres étaient également recueillis en cas de

décision de transfert en réanimation. Le recueil de ces paramètres de surveillance clinique a été appelé « observations » dans ce travail.

#### 1.6.4- Traitement et saisie des données recueillies

Le service de référence était l'équipe d'accueil 2694 du Pr Alain Duhamel, au sein duquel ont été réalisés :

- la construction du masque de saisie du tableau de bord (Capture System)
- la centralisation des données informatiques
- le contrôle de la qualité des données saisies (Capture System)
- la sécurisation de la gestion de la base de données.

Les données cliniques (caractéristiques générales et feuille de surveillance) étaient saisies en utilisant un formulaire électronique (E-CRF). Le formulaire a été développé en utilisant un logiciel professionnel et était le reflet exact du cahier d'observation (caractéristiques générales de chaque patient et des différentes feuilles de surveillance). Le masque était personnalisé afin d'assurer la qualité de la saisie avec différentes règles de saisie : types de données (format pré spécifié, code spécifique pour les données manquantes, variables qualitatives présentées sous la forme de listes déroulantes ou de boutons), propriété de champ : taille du champ, bornes de valeurs possibles pour les données numériques ou les dates, règle de validation de champ, règle de validation d'enregistrement. Ce travail de modélisation des données, schéma conceptuel, développement de la base de données, procédures de transfert des données était réalisé avec le logiciel Capture System. Des règles de cohérence des données étaient développées et des analyses descriptives étaient réalisées pour la validation des données : box plot (boîtes à moustaches) pour les données numériques, identification du taux de données manquantes pour chaque variable, détection des individus aberrants ou exceptionnels, variable par variable, détection des incohérences ou des individus aberrants en multivarié par analyse de classification automatique. Des rapports étaient édités comprenant le taux de données manquantes, les incohérences, les valeurs aberrantes. Des "queries" étaient envoyés aux investigateurs et des corrections étaient effectuées après vérification.

Une convention de prestation était passée avec le Centre François Baclesse de Caen, organisme gestionnaire du Centre de Traitement des Données (CTD) du Cancéropôle Nord-Ouest, plate-forme labellisée par la Ligue Nationale Contre le Cancer pour :

- l'hébergement des données,
- la création des accès au logiciel de saisie,
- la maintenance.

### **1.7- Calcul du nombre de sujet nécessaire**

Le nombre de sujets nécessaires a été calculé par la méthode basée sur l'article de Hanley JA and McNeil BJ<sup>46</sup> qui permet de calculer le nombre de sujets nécessaires pour montrer qu'une AUC est au moins égale à une valeur minimale fixée. Pour la validation des SAP dans notre étude, nous avons fixé la valeur minimale de l'AUC à 0,75 correspondant à un pouvoir diagnostique correct.

Dans la littérature, le pourcentage d'appels d'un médecin par l'infirmière auprès d'un enfant hospitalisé en secteur classique était d'environ 10%<sup>22</sup>. En se basant sur un taux d'appel du médecin par l'infirmière de 10% dans notre étude, en considérant une AUC attendue de 0,80, un risque de première espèce de 0,05 et un test unilatéral, il fallait inclure 2000 patients pour obtenir une AUC >0,75 avec une puissance de 80% et 2800 patients avec une puissance de 90%. Nous avons estimé le nombre de sujet nécessaire pour notre étude à 2500 patients, ce qui correspondait au recrutement attendu des centres participants sur un an (estimation à partir des données de 2009 et 2010 dans chacun des centres participants), permettant d'observer 250 évènements et correspondant à une puissance de 87%.

### **1.8- Analyses statistiques**

Lorsqu'un ou deux des paramètres cliniques manquaient, le score était calculé en considérant les valeurs manquantes comme normales. Les scores avec plus de deux paramètres manquants n'étaient pas analysés. Une analyse de sensibilité a été effectuée pour déterminer l'impact de l'imputation d'une valeur normale par défaut par rapport à une analyse de cas complets (ces résultats sont présentés dans le tableau A1, Annexe 7). La reproductibilité interobservateur (infirmière-médecin) a été évalué en calculant le coefficient de corrélation intraclasse pour chaque score.

La cohorte a été divisée en cohortes de dérivation (les premiers 70% inclus) et de validation (les 30% restants). Dans la cohorte de dérivation, un modèle de régression logistique avec un effet aléatoire (modèle linéaire mixte généralisé) a été utilisé pour prendre en compte plusieurs observations par sujet. Dans ce modèle, les effets fixes

étaient le temps  $t$ , le score à l'admission ( $H_0$ ) et le score au temps  $t$ . Ce modèle a permis d'estimer la probabilité prédite d'appel du médecin par l'infirmière à chaque instant  $t$ . Le pouvoir discriminant du modèle associée aux probabilités prédites a été estimé par l'AUC. Le rapport de cotes (Odds ratio des anglo-saxons (OR)) de chaque score, pour l'appel du médecin correspondant à une augmentation d'un point de score, ajusté sur le score à l'admission et le temps, a été calculé avec leur intervalle de confiance à 95%. Pour chaque score, nous avons calculé le seuil de la probabilité prédite, en maximisant la spécificité pour une sensibilité fixée à 90%, ce qui a permis d'établir une règle de prédiction pour l'appel du médecin par l'infirmière.

Dans la cohorte de validation, les AUC ont été calculées à l'aide des coefficients des effets fixes estimés sur la cohorte de dérivation, comme indiqué par Foulkes et *al.*<sup>47</sup>. La stabilité des règles de prédiction a été évaluée en calculant la sensibilité et la spécificité correspondant aux seuils déterminés dans la cohorte de dérivation. Les analyses ont été effectuées en utilisant les observations répétées du score de 8h à 24 heures (durée médiane du séjour), de 8h à 36 heures, de 8h à 48 heures et de 8h à 6 jours (6 jours correspondant au 95<sup>ème</sup> percentile de la durée du séjour en USC). Toutes les analyses ont été ajustées sur la valeur du score à l'admission ( $H_0$ ).

Toutes les analyses ont été effectuées à l'aide d'un test bilatéral avec un niveau alpha de 0,05. Les analyses statistiques ont été effectuées à l'aide du logiciel SAS version 9.4 (SAS Institute, Cary, NC, États-Unis).

Nous avons définis des critères d'exclusion de l'analyse statistique:

Les patients qui n'avaient pas de scores à  $H_0$  ou pour lesquels aucun score n'a pu être calculé après  $H_0$  ont été exclus et les observations enregistrées après 6 jours ont été exclues de l'analyse.

### **1.9- Considérations éthiques**

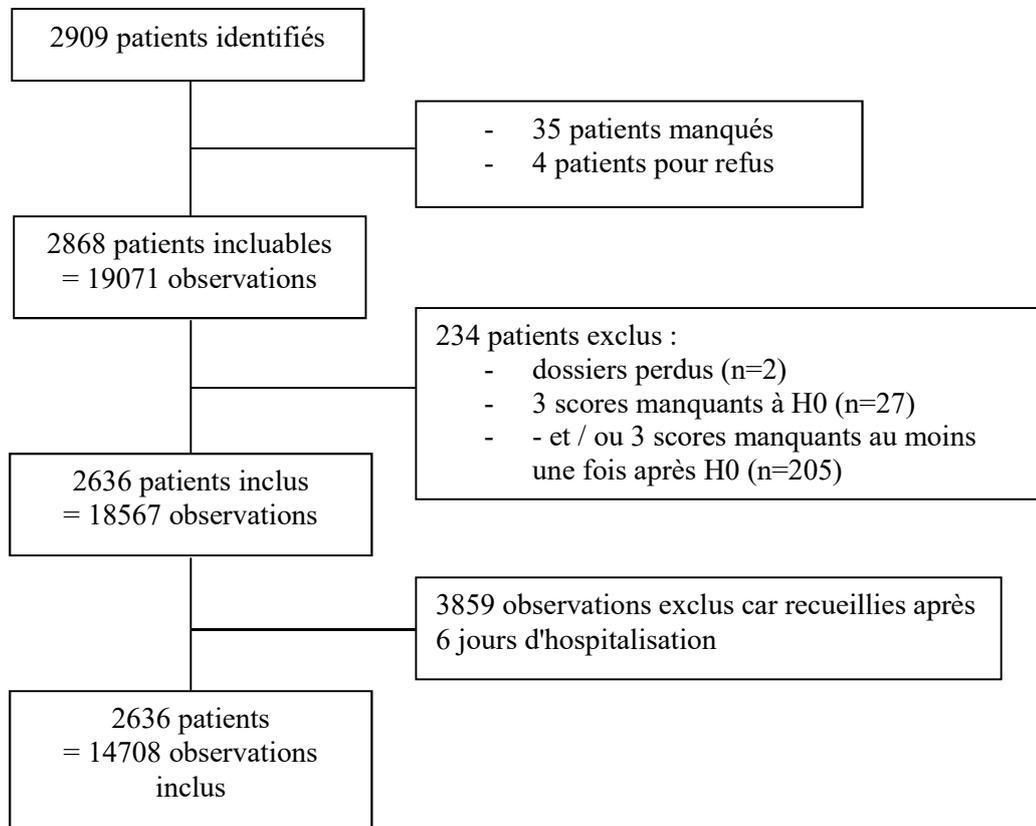
Aucune intervention diagnostique ou thérapeutique supplémentaire, aucune modalité particulière de surveillance n'était réalisée aux fins de l'étude. Il s'agissait d'une recherche non interventionnelle (mentionnée au 1<sup>o</sup> de l'article L. 1121-1 et à l'article R. 1121-2 du code de la santé publique), strictement observationnelle. La déclaration à la Commission Nationale d'Informatique et Liberté (CNIL) a été effectuée et un accord favorable a été émis le 17 décembre 2012 (DR-2012-594). Le comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé (CCTIRS) a émis un avis favorable en date du 03/05/2012 (N° 12.253). Cette étude a

également été déclarée au registre international des essais cliniques (ClinicalTrials.gov, NCT02304341). De plus, l'étude a été soumise à la commission d'éthique de la société de réanimation de langue française (SRLF) et a obtenu un avis favorable le 02 février 2012 (CE SRLF 12-351).

Une note d'information était affichée dans chaque centre participant et était remise en main propre à tous les parents (Annexe 8) pour les aviser de la réalisation d'une étude observationnelle à propos de l'évaluation de la gravité des enfants hospitalisés en USC pédiatriques, et leur rappeler leur droit d'accès et de rectification des données les concernant, ainsi que leur droit d'opposition au traitement de celles-ci, conformément à la loi du 6 janvier 1978. D'autre part, pour les enfants en âge de lecture, une note d'information adaptée leur a été remise (Annexe 9). Une traduction en Arabe des documents d'information était disponible pour répondre aux populations présentes dans la région de l'étude. En cas d'opposition des parents à la participation à l'étude, un soignant a notifié le refus sur un formulaire (Annexe 10).

## 2- Résultats

Sur les 2 909 patients évalués pour l'éligibilité, 2 636 ont été inclus. Le flow chart est présenté à la figure 1 et les caractéristiques des patients sont présentés tableau 1.



**Figure 1: Flow chart des inclusions**

La comparaison des caractéristiques générales des patients inclus et des patients exclus (234 patients versus 2636) est présentée au tableau 1. Parmi les patients exclus, 88% (205/234) l'étaient parce qu'aucun des trois scores n'avaient pu être calculés au moins une fois après H0, ce qui implique une très courte durée de séjour (<8 heures). Ces patients étaient plus âgés (42 mois [7-162] versus 28 mois [5-98]), plus souvent admis directement depuis le domicile (20,9% versus 12,4%), avait une plus faible incidence de défaillance respiratoire (34,6% versus 45,2%) à l'admission et un taux d'infection plus faible (41,0% versus 55%) que les patients inclus ( $p < 0,05$ ). Ils avaient également une médiane de score à l'admission plus faible pour les 3 scores ( $p < 0,05$ ), ce qui est cohérent avec une durée de séjour en USC pédiatrique très courte inférieures à 8h.

**Tableau 1: Caractéristiques générales des patients hospitalisés en USC pédiatriques**

Caractéristiques générales	Total, n=2870	Patients inclus n=2636	Patients Exclus* n=234	p
Sex ratio	1,26	1,26	1,32	ns
Age median (mois) [IQR]	29 [5-103]	28 [5-98]	42 [7-162]	0,01
Durée de séjour médaine (jours) [IQR]	1,0 [1,0-3,0]	1,5 [0,8-2,7]	0,2 [0,1-0,3]	< 10 <sup>-4</sup>
Comorbidités, n(%)	945 (32,9)	876 (33,2)	69 (29,5)	ns
Provenance, n(%)				0,001
Urgences	2006 (69,9)	1851 (70,2)	155 (66,2)	
Pédiatrie	386 (10,0)	361 (13,7)	25 (10,7)	
Bloc	71 (2,5)	68 (2,6)	3 (1,3)	
Réanimation	25(0,9)	25 (0,9)	0 (0,0)	
Domicile	377(13,1)	328 (12,4)	49 (20,9)	
Type de séjour, n(%)				ns
Medical	2714 (94,6)	2490 (96,5)	224 (95,7)	
Chirurgical	119 (4,1)	113 (4,3)	6 (2,6)	
Défaillance aiguë principale, n(%)				0,002
Respiratoire	1272 (44,3)	1191(45,2)	81(34,6)	
Neurologique	627 (22,8)	580(22,0)	47(20,0)	
Cardiovasculaire	218 (7,6)	198(7,5)	20(8,5)	
Gastrointestinal	222 (7,7)	202(7,7)	20(8,5)	
Autres	529 (18,4)	465 (17,6)	64 (27,3)	
Etiologies, n(%)				< 10 <sup>-4</sup>
Infectieuse				
Toxique	1546 (54,9)	1450 (55,0)	96 (41,0)	
Congenital	201 (7,0)	176 (6,7)	25 (10,7)	
Traumatique	136 (4,7)	127 (4,8)	9 (3,8)	
Oncologique	126 (4,4)	114 (4,3)	12 (5,1)	
Autres	65 (2,3)	42 (1,6)	23 (9,8)	
Destination n(%)	794 (27,7)	727 (27,6)	67 (28,6)	
Domicile	1024 (35,7)	937 (35,5)	87 (37,2)	
Pédiatrie	1586 (55,3)	1487 (56,4)	99 (42,3)	
Autres	246 (8,6)	201 (7,6)	45 (19,2)	
Décès, n	1	1	0	
Score median à l'admission [IQR]				
PAWS	2 [0-4]	3 [2-4]	2 [1-4]	0,001
PEWS	0 [0-2]	1 [0-3]	0 [0-3]	0,005
Bedside PEWS	3 [0-5]	3 [2-6]	3 [1-5]	0,004

ns: non significatif, IQR: interquartile, PAWS: Pediatric Advanced Warning Score; PEWS: Pediatric Early Warning Score; Bedside PEWS: Bedside Pediatric Early Warning System

\*234 patients exclus : dossier perdu (n=2), au moins un score manquant à H0 (n=27), au moins un score non calculable après H0 (n=205).

La reproductibilité des trois scores était bonne : les valeurs des coefficients de corrélation intra-classe (IC 95%) pour les scores PAWS, PEWS et Bedside PEWS étaient de 0,807 (0,802-0,812) (n = 372 observations), 0,857 (0,853-0,861) (n = 375) et 0,806 (0,801-0,811) (n = 367), respectivement.

Sur les 19071 observations recueillies, 14 708 étaient incluses dans l'analyse permettant de calculer 12 668 (86,1%) PAWS, 11 756 (79,9%) PEWS et 12 191 (82,9%) Bedside PEWS scores complets et 1 710 (11,6%) PAWS, 2 916 (19,8%) et 2 078 (14,1%) Bedside PEWS

supplémentaires, après avoir considéré comme normal un ou deux paramètres manquants (figure 1). L'appel du médecin par l'infirmière était renseigné « oui » pour 1 064 observations (7%). Les scores médians lorsqu'il y avait « l'appel au médecin par l'infirmière » étaient de 3 [2-5] pour le PAWS, de 2 [0-3] pour le PEWS et de 5 [2-8] pour le Bedside PEWS.

Les performances des trois scores pour prédire l'appel du médecin par infirmière sur la cohorte de dérivation sont présentées dans le tableau 2. Ces résultats portent sur les scores complets et les scores imputés avec une valeur normale par défaut à la jusqu'à deux items manquants. Les résultats de l'analyse de sensibilité montraient des AUC comparables (Tableau A1, Annexe 7) que les scores imputés soient ajoutés ou non à l'analyse des scores complets. L'AUC des trois scores se situait entre 0,87 ( $\leq 24$  heures) et 0,91 ( $\leq 6$  jours). Les équations prédictives de l'appel du médecin par l'infirmière, tenant compte du temps  $t$ , du score au temps  $t$  et du score à l'admission ( $H_0$ ) sont présentées en légende du tableau 2.

**Tableau 2: Performances du PAWS, PEWS et du Bedside PEWS sur la cohorte de dérivation sur les 1<sup>ères</sup> 24 heures, les 1<sup>ères</sup> 36 heures, les 1<sup>ères</sup> 48 heures, et les 6 1<sup>ers</sup> jours**

Scores	AUC (95% IC)	Seuil de probabilités prédites (%) <sup>a</sup>	Spécificité (%)	VPP (%)	VPN (%)
<b>≤ 24 heures [n=4,693 observations; Appel du médecin = 494 (11%) observations]</b>					
PAWS <sup>b</sup>	0,87 (0,85-0,89)	5,9	49	17	98
PEWS <sup>c</sup>	0,87 (0,86-0,89)	6,0	48	17	98
Bedside PEWS <sup>d</sup>	0,87 (0,85-0,89)	5,2	48	16	98
<b>≤ 36 heures [n=5,892 observations; Appel du médecin = 559 (9%) observations]</b>					
PAWS <sup>b</sup>	0,90 (0,88-0,91)	7,3	64	20	98
PEWS <sup>c</sup>	0,89 (0,88-0,91)	6,7	55	17	98
Bedside PEWS <sup>d</sup>	0,89 (0,87-0,91)	6,2	58	17	98
<b>≤ 48 heures [n=7,196 observations; Appel du médecin = 629 (9%) observations]</b>					
PAWS <sup>b</sup>	0,91 (0,90-0,92)	7,5	69	22	99
PEWS <sup>c</sup>	0,91 (0,89-0,92)	7,4	71	23	99
Bedside PEWS <sup>d</sup>	0,91(0,89-0,92)	6,7	70	21	99
<b>≤ 6 jours [n=10,152 observations; Appel du médecin = 749 (7%) observations]</b>					
PAWS <sup>b</sup>	0,91 (0,90-0,92)	7,1	77	23	99
PEWS <sup>c</sup>	0,91 (0,90-0,92)	7,1	76	23	99
Bedside PEWS <sup>d</sup>	0,91 (0,90-0,92)	6,1	75	21	99

<sup>a</sup> Probabilité maximisant la spécificité pour une sensibilité fixée à 90%: probabilité calculée à partir du modèle linéaire mixte généralisé incluant les effets suivants: le score à l'instant t, le score à l'admission et le temps t.

<sup>b</sup> équation prédictive de l'appel du médecin =  $-2,867 + 0,315*(\text{score au temps t}) - 0,072*(\text{score à l'admission}) - 0,016*(\text{temps t})$

<sup>c</sup> équation prédictive de l'appel du médecin =  $-2,630 + 0,448*(\text{score au temps t}) - 0,101*(\text{score à l'admission}) - 0,017*(\text{temps t})$

<sup>d</sup> équation prédictive de l'appel du médecin =  $-3,033 + 0,263*(\text{score au temps t}) - 0,058*(\text{score à l'admission}) - 0,015*(\text{temps t})$

Abbreviations: AUC: aire sous la courbe; IC: intervalle de confiance; PAWS: Pediatric Advanced Warning Score; PEWS: Pediatric Early Warning Score; Bedside PEWS: Bedside Pediatric Early Warning System; VPP: valeur predictive positive; VPN: valeur predictive négative.

Deux exemples d'utilisation des équations prédictives pour le score PAWS sont présentés ci-dessous :

Exemple 1:

Si le PAWS à l'admission =3, temps=48 heures et le PAWS (H48)=4

Equation predictive pour le PAWS =  $-2,867 + 0,315*(\text{score au temps } t) - 0,072*(\text{score à l'admission}) - 0,016*(\text{temps } t)$

Logit=  $-2,867 + 0,315*(4) - 0,072*(3) - 0,016*(48) = -2,591$

Probabilité d'appel du médecin =  $1/(1+e^{-\text{Logit}}) = 7,0\%$

Cette probabilité est inférieure à 7,5% (correspondant au seuil de probabilité pour le PAWS sur la période admission-H48, voir tableau 2): donc le modèle prédit : pas d'appel du médecin par l'infirmière.

Exemple 2:

Si le PAWS à l'admission=0, temps=48 heures, PAWS (H48)=4

Equation pour le PAWS =  $-2,867 + 0,315*(4) - 0,072*(0) - 0,016*(48) = -2,375 (= \text{logit})$

$p = 1/(1+e^{-\text{logit}}) = 8,5\%$

$p > 7,5\%$  (correspondant au seuil de probabilité pour le PAWS sur la période admission-H48, voir tableau 2) : donc le modèle prédit que l'infirmière doit appeler le médecin.

Les spécificités des trois scores (PAWS, PEWS et Bedside PEWS ) étaient respectivement de 48% et 49% si on tient compte des observations  $\leq 24$  heures et augmentées à 75 et 76% si on tient compte des observations  $\leq 6$  jours (tableau 2). Les valeurs prédictives négatives étaient excellentes (98 et 99%) pour tous les scores et à toutes les périodes de temps (tableau 2).

Chaque augmentation d'un point de PAWS, de PEWS et de Bedside PEWS augmentait de manière significative le risque d'appel du médecin (OR [IC 95%]: 1,37 [1,30-1,44], 1,57 [1,47-1,67] et 1,30 [1,26- 1,35], respectivement).

La règle de prédiction construite à partir de la cohorte de dérivation (si la probabilité prédite était supérieure au seuil) était appliquée à la cohorte de validation (tableau 3).

**Tableau 3: Performances du PAWS, PEWS et du Bedside PEWS sur la cohorte de validation sur les 1<sup>ères</sup> 24 heures, les 1<sup>ères</sup> 36 heures, les 1<sup>ères</sup> 48 heures, et les 6 1<sup>ers</sup> jours**

	AUC (95% IC)	Sensitivité (%)	Spécificité (%)	VPP (%)	VPN (%)
<b>≤ 24 heures [n=2,028 observations; Appel du médecin = 219 (11%) observations]</b>					
PAWS	0.74 (0.70-0.78)	76	50	15	95
PEWS	0.74 (0.70-0.78)	77	48	15	95
Bedside PEWS	0.76 (0.72-0.80)	81	46	15	95
<b>≤ 36 heures [n=2,550 observations; Appel du médecin = 245 (10%) observations]</b>					
PAWS	0.74 (0.70-0.77)	72	60	16	95
PEWS	0.75 (0.71-0.78)	78	53	15	96
Bedside PEWS	0.75 (0.72-0.79)	79	53	15	96
<b>≤ 48 heures [n=3,156 observations; Appel du médecin = 267 (8%) observations]</b>					
PAWS	0.73 (0.70-0.77)	64	71	17	96
PEWS	0.75 (0.72-0.79)	73	63	15	96
Bedside PEWS	0.75 (0.72-0.79)	71	65	15	96
<b>≤ 6 jours [n=4,556 observations; Appel du médecin = 315 (7%) observations]</b>					
PAWS	0.71 (0.68-0.74)	55	76	14	96
PEWS	0.75 (0.72-0.78)	58	75	15	96
Bedside PEWS	0.75 (0.72-0.78)	66	69	13	96

Les performances des scores ont été calculées à partir de l'équation prédictive dérivée de l'échantillon de dérivation; la sensibilité et la spécificité ont été calculées à l'aide des seuils mesurés sur l'échantillon de dérivation

Abbreviations: AUC: Aire sous la courbe; IC: intervalle de confiance; PAWS: Pediatric Advanced Warning Score; PEWS: Pediatric Early Warning Score; Bedside PEWS: Bedside Pediatric Early Warning System, VPP: valeur prédictive positive; VPN: valeur prédictive négative.

Sur l'échantillon de validation, l'AUC des trois scores se situait entre 0,71 (PAWS ≤ 6 jours) et 0,76 (Bedside PEWS ≤ 24 heures). Les sensibilités des trois scores étaient maximales à H24 (76 à 81%) et diminuaient jusqu'au jour 6 (55% à 66%), tandis que la fréquence des appels au médecin diminuait également de 11 à 7%. Les valeurs prédictives négatives étaient excellentes (95 et 96%) pour tous les scores et sur toutes les périodes de temps.

### **3- Discussion**

Cette étude est la première à évaluer l'utilisation des SAP dans les USC pédiatriques pour détecter la détérioration des enfants. La discrimination estimée par les AUC des trois scores pour prédire l'appel du médecin par l'infirmière était bonne à excellente (fourchette: 0,87 à 0,91) sur la cohorte de dérivation et modérée (fourchette: 0,71 à 0,76) sur la cohorte de validation. Les équations prédictives d'appel du médecin, tenant compte du score au temps t, du score à l'admission et du temps t, sont disponibles. Les trois SAP avaient une bonne reproductibilité interobservateur

#### **3.1- Choix du critère de jugement**

Quatre revues sur les SAP pédiatriques ont été publiées<sup>15,17,34,48</sup>. Pour développer ou valider les SAP pédiatriques, un gold standard établissant la détérioration clinique est nécessaire, bien qu'il n'y ait pas de consensus sur ce gold standard 3. Plusieurs critères de jugement ont été utilisés: décès<sup>18</sup>, arrêt cardiaque ou code bleu<sup>19-21</sup>, transfert vers une réanimation<sup>21-27,49,50</sup>, appel pour une assistance médicale urgente ou activation du système de réaction rapide (SRR)<sup>28-31</sup> et la durée du séjour à l'hôpital<sup>28</sup>. Cependant, le but de ces SAP est d'identifier la détérioration avant un arrêt respiratoire ou cardiaque ou un transfert en réanimation. De plus, la plupart de ces critères sont des événements rares, ce qui affecte la méthodologie utilisée pour la validation des SAP.

Dans notre étude, nous avons le choix entre deux critères de jugement :

- transfert en réanimation
- appel du médecin en cas de détérioration détectée par l'infirmière.

La limite du critère de jugement « transfert en réanimation » est qu'il ne prenait en compte que les détériorations cliniques les plus graves, nécessitant le transfert de l'enfant. Or, un SAP devrait pouvoir détecter les patients avant que leur état ne nécessite un transfert en réanimation, de manière à ce que le médecin puisse intervenir et éviter cet événement.

Nous avons retenu comme critère de jugement l'appel du médecin par l'infirmière.

Ce critère présente également des limites car il est influencé par de nombreux facteurs: expérience de l'infirmière, connaissance du patient, relation entre l'infirmière et le médecin, facilité d'appel et charge de travail. L'appel du médecin par l'infirmière peut être un facteur de confusion, car la collecte de paramètres cliniques au moment de la «détérioration» dépend du jugement clinique subjectif de l'infirmière. Il n'a pas été possible de vérifier que l'appel du médecin par l'infirmière était toujours justifié. Bonafide et *al.*, dans une étude qualitative sur l'utilisation des SAP par le personnel infirmier et les médecins, ont suggéré que la combinaison de ces scores et du jugement clinique pourrait constituer un meilleur système de détection de la détérioration<sup>51</sup>. Jensen et *al.*, dans une étude s'intéressant aux facteurs qui pourraient compromettre l'utilisation des SAP en pratique clinique, signalaient aussi le manque de jugement clinique<sup>52</sup>.

De plus, le critère de jugement « transfert en réanimation pédiatrique » était un événement trop rare dans notre étude pour être utilisé dans une méthodologie prospective incluant tous les patients consécutifs. En effet, on ne comptait que 85 scores recueillis lors d'une décision de transfert en réanimation sur un total de 14708 scores calculables.

### **3-2- Choix d'un modèle linéaire mixte généralisé pour mesures répétées**

Dans notre étude, nous avons à analyser des données répétées. Pour chaque patient inclus, nous disposons d'observations à différents instants, répétées dans le temps (à l'admission, toutes les 8 heures, et en cas de détérioration). La difficulté dans le traitement statistique de ces données est liée au fait que les observations réalisées chez un même individu, au cours du temps, ne peuvent pas être considérées indépendantes entre elles. En pratique courante, les SAP sont répétés à intervalles réguliers et il semblait pertinent de prendre en compte l'ensemble des SAP chez un même patient.

Nous avons choisi d'utiliser un modèle mixte pour données répétées, afin de considérer, en plus des facteurs à effets fixes, des effets aléatoires associés aux individus. La description détaillée du modèle est présentée Annexe 11. De nombreux SAP ont été développés de manière rétrospective et leur validité a été évaluée en utilisant souvent une méthodologie de type cas / contrôle avec un nombre de cas jamais supérieur à 120 et un seul score par patient<sup>20,21,24,27,49,53,54</sup>.

### **3-3- Proposition d'utiliser les scores d'alerte précoce dans une équation prédictive d'appel du médecin**

En tenant compte de l'ensemble des mesures de chaque patient, un seuil unique de score n'a pu être identifié en raison de mesures répétées chez le même patient.

Dans la littérature, les seuils proposés pour les SAP sont variables et ne peuvent être comparés entre eux<sup>34</sup>. Dans des études cas/témoin, utilisant comme critère de jugement l'arrêt cardiaque, l'activation du système de réaction rapide ou l'admission en réanimation, les seuils proposés pour le Bedside PEWS sont élevés  $\geq 7$  et  $\geq 8$  selon les critères choisis<sup>20,23-25,53</sup>. Pour différents scores PEWS modifiés, les seuils proposés étaient compris entre  $\geq 2$  et  $\geq 5$ <sup>19,20,22,26,49,54</sup> et pour PAWS  $\geq 3$ <sup>43</sup>.

Dans deux études prospectives, utilisant comme critère de jugement l'admission en réanimation, un seuil  $\geq 1$  pour le PEWS<sup>55</sup> et  $\geq 3$  pour le PAWS<sup>56</sup> ont été proposés. Seiger *et al.*, ont analysé la validité de différents SAP chez 17 943 enfants et observé que le seuil optimal pour prédire l'admission en réanimation était faible : seuil à 1 pour tous sauf pour le BedsidePEWS pour lequel le seuil était de 3<sup>57</sup>.

Gerry *et al.* expliquaient que les méthodologies utilisées pour le développement ou la validation des SAP pourraient être responsables du manque de fiabilité et de la mauvaise prédiction du risque par ces scores<sup>32</sup>. La raison principale, selon Chapman, pouvant expliquer le manque de preuve d'efficacité des SAP est le taux faible d'événement signant la détérioration<sup>58</sup> dans les études.

Certains auteurs ont proposé une méthodologie plus robuste. En pédiatrie, Jensen *et al.*, ont mené un essai multicentrique contrôlé et randomisé comparant deux modèles différents de SAP pour prédire une détérioration nécessitant un transfert vers un niveau de soins plus élevé<sup>59</sup>. Dans cette étude, malgré un grand nombre de patients inclus ( $n = 16213$ ), seulement 22 patients ont été transférés vers un niveau de soins plus élevé. Aucune différence significative entre les deux scores (BedsidePEWS et Central Denmark Region :CDR PEWS) n'était identifiée, mais le CDR PEWS semblait plus acceptable pour

le personnel<sup>59</sup>. Parshuram et *al.*, ont évalué l'effet du Bedside PEWS sur la mortalité hospitalière toutes causes confondues, l'admission tardive en réanimation, l'arrêt cardiaque et le recours aux équipes de réanimation, dans un essai clinique randomisé, comprenant 144 539 patients (21 hôpitaux dans sept pays). Dans cette étude, le score Bedside PEWS était appliqué à tous les malades, avec mise en place d'un programme d'éducation des équipes. En fonction de la valeur du Bedside PEWS relevé par l'infirmière lors de sa surveillance, un programme structuré d'escalade ou de désescalade de prise en charge était préconisé. En dépit de ce très grand nombre de patients, pour compenser le faible nombre d'événements, l'utilisation du Bedside PEWS n'a pas diminué significativement la mortalité toutes causes confondues, par rapport aux soins habituels<sup>60</sup>. Selon Chapman et *al.*, le décès n'est pas le bon critère de jugement pour l'évaluation des SAP, tout comme le transfert en réanimation, la réanimation cardio-pulmonaire, l'intubation trachéale qui sont des critères témoignant plutôt d'une détection tardive de la détérioration, empêchant toute action préventive<sup>58</sup>. De ce fait, les seuils proposés dans la littérature sont probablement trop élevés pour permettre une intervention préventive.

Nous proposons dans notre étude une approche différente par le choix du critère de jugement : l'appel du médecin par l'infirmière qui permet cette action préventive et par la prise en compte de la répétition des mesures du score chez un même individu.

### **3-4- Evolutions de la performance des scores au cours de l'hospitalisation**

Dans notre étude, les AUC des trois scores étaient bonnes à excellentes (0,87-0,91) pour la cohorte de dérivation et modérées (0,71-0,76) pour la cohorte de validation.

Toutes les analyses étaient effectuées sur différentes périodes de temps pour évaluer l'évolution de la performance des SAP au cours de l'hospitalisation. Pour cela, les analyses étaient effectuées en utilisant les observations de l'admission à 24 heures (correspondant à la durée médiane du séjour), de l'admission à 36 heures, de l'admission à 48 heures et de l'admission jusqu'à 6 jours (6 jours correspondant au 95ème percentile de la durée du séjour).

Sur la cohorte de validation, la sensibilité diminuait (de 76-81% à 55-66%) et la spécificité augmentait avec le temps pour les trois scores (de 46-50% à 69-76%).

Mulherin et *al.*, indiquait que les variations de sensibilité et de spécificité en fonction de sous-groupes n'étaient pas un biais mais représentait une information cliniquement pertinente à identifier<sup>61</sup>. Ces auteurs suggéraient de remplacer le terme «biais spectral » par« effet de spectre » reflétant l'hétérogénéité de la performance des scores lorsqu'ils

sont appliqués à différents sous-groupes. Dans notre étude, nous avons inclus différents sous-groupes d'observations selon la durée d'hospitalisation depuis l'admission. Ces trois scores semblaient mieux détecter la détérioration au début du séjour et leurs performances semblaient diminuer avec le temps. Cela pourrait être lié à la stabilisation des patients, nécessitant ainsi de moins en moins d'appels du médecin au fur et à mesure du séjour.

### **3-5- Reproductibilité des scores d'alerte précoce**

La reproductibilité des trois scores était bonne (coefficients de corrélation intra-classes : 0,81 à 0,86). Peu d'études ont analysé la reproductibilité inter-observateur de ces SAP: Chaiyakulsil *et al.* ont signalé une bonne fiabilité inter-évaluateur ( $\kappa = 0,75$ ) pour le PAWS<sup>56</sup> et Gold *et al.* ont signalé une excellente fiabilité inter-évaluateur (coefficient intra-classe = 0,91) pour le PEWS<sup>55</sup>. Le choix du jour et l'heure de la collecte simultanée de données par les équipes pour l'évaluation de la reproductibilité peut constituer un biais si les équipes choisissent des moments où les unités sont calmes pour ce recueil. Le recueil des différents paramètres physiologiques a été effectué par l'infirmière responsable de l'enfant et par un deuxième observateur : le médecin en charge de l'enfant. L'horaire de collecte a été choisi par l'équipe de chaque centre et devait rester le même tout au long de l'étude (par exemple tous les lundis à 11h00). Une des qualités indispensables pour l'efficacité d'un SAP est la facilité avec laquelle il est possible de collecter les données, de calculer le score et d'interpréter le score. Nous n'avons pas pu évaluer tous ces éléments puisque le score était calculé à postériori. Avec cette étude de reproductibilité, nous voulions vérifier l'adéquation entre l'infirmière et le médecin pour des éléments subjectifs pouvant varier d'un observateur à l'autre, tels que le temps de recoloration capillaire, l'effort respiratoire ou le comportement de l'enfant.

### **3-6- Gestion des données manquantes**

Pour le traitement des données manquantes nécessaire au calcul des scores, plusieurs possibilités étaient envisagées.

La première possibilité était d'analyser uniquement les scores complets, c'est-à-dire ceux pour lesquels aucun item (paramètres physiologiques cliniques) n'était manquant. Cela impliquait une perte de donnée et un biais de sélection potentiel. Les données manquantes pour chaque score sont présentées dans le tableau A2, Annexe 12.

La deuxième possibilité était de considérer les item manquants comme normaux et de calculer le score en attribuant à ces item manquants « 0 point », à condition que le nombre d'items manquants soit strictement inférieur à 3 sur le total des valeurs nécessaires au calcul de chacun des 3 scores. Le détail de données imputées est présenté dans le tableau A2, Annexe 12. Ainsi, lorsqu'un ou deux des paramètres cliniques manquaient, le score était donc calculé en considérant les valeurs manquantes comme normales. Les scores avec plus de deux paramètres manquants étaient exclus de l'analyse. Une analyse de sensibilité était effectuée pour déterminer l'impact de l'imputation par une valeur normale par rapport à une analyse de cas complet (Tableau A1, Annexe 12). Il n'y avait aucune différence entre les AUC. La méthode d'imputation par une valeur normale jusqu'à 2 items manquants par score était donc retenue.

Une autre possibilité aurait été de réaliser des imputations multiple. Nous n'avons pas opté pour cette méthode en raison de sa complexité dans le contexte de notre étude (mesures répétées, modèle mixte avec effets aléatoires).

La gestion des données manquantes n'est pas souvent explicitée dans les études sur les SAP. Parshuram et *al.*<sup>24</sup> ont pris la donnée enregistrée la plus récente pour les items manquants<sup>24</sup>. Certains auteurs ont attribué des valeurs normales aux items manquants, quel que soit le nombre d'items manquants<sup>20,21,43</sup>. D'autres ont utilisés un modèle d'imputation multiple<sup>56,57</sup>.

### **3-7- L'effet centre**

Les sept USC de l'étude étaient situées dans des hôpitaux généraux d'une même région. Elles n'étaient pas rattachées à un service de réanimation pédiatrique. Elles comprenaient chacune 4 à 6 lits. Le rapport infirmière/patient était de 1/4.

Nous avons étudié les caractéristiques des patients entre les sept centres. Il existait des différences sur les valeurs d'âges des patients, de durée de séjour d'appel du médecin par l'infirmière ou de durée de séjour (tableau A3, Annexe 13). Cependant, lorsque nous avons ajusté sur le centre, le pouvoir discriminant évalué par l'AUC était similaire pour les trois scores (tableau A4, Annexe 13). L'effet centre n'a pas été ajouté en tant qu'effet aléatoire dans le modèle car il ne permettait pas l'obtention d'une règle globale de prédiction d'appel du médecin par l'infirmière.

### **3-8- Représentativité des patients d'une USC non attachée à une réanimation pédiatrique**

Les patients inclus dans notre étude étaient des patients admis en USC pédiatriques de centre hospitalier généraux, non attachées à une réanimation pédiatrique. Ces patients pourraient être différents (peut-être moins sévères) que des patients hospitalisés dans les USC pédiatriques attachées à une réanimation. Comme indiqué par Plate *et al.*, il existe une grande diversité de ces USC<sup>62</sup>. Dans notre population, le taux de transfert vers la réanimation était de 3%. Ce taux est proche de celui rapporté par Seiger *et al.* dans leur étude faite dans un service d'urgence (2%)<sup>57</sup>. Les patients de notre étude étaient peut-être davantage comparables aux patients de service de soins courants ou des urgences pédiatriques qu'aux patients de réanimation pédiatrique.

### **3-9- Quelle conclusion de notre article et quel score d'alerte précoce choisir en pratique?**

Dans notre étude, aucun des trois scores ne semblait être meilleur que les autres pour détecter la détérioration. Cependant, le PAWS peut être préféré pour une évaluation initiale rapide en soins d'urgence ou dans les services de soins courants parce que la mesure de la pression artérielle n'est pas requise. Comme nous l'avons dit précédemment, le Bedside PEWS présente l'avantage de prendre en compte la saturation en oxygène et les apports en oxygène, contrairement au PAWS, qui ne considère que la saturation en oxygène. L'inconvénient du Bedside PEWS est qu'il n'inclut aucune évaluation neurologique, contrairement aux 2 autres scores. Le PEWS est divisé en trois catégories: respiratoire, circulatoire et neurologique, correspondant aux différents types de défaillance d'organes et, permet donc possiblement une meilleure visibilité, une meilleure synthèse de la situation.

Sambecq et *al.* dans une enquête transversale, ont montré que 45 SAP différents étaient utilisés dans les hôpitaux néerlandais. Différents auteurs ont signalé que l'utilisation large de ces SAP en pratique quotidienne, alors qu'aucun consensus n'existe sur le SAP à utiliser, ne semble reposer sur aucune méthodologie robuste de validation et qu'aucun seuil pour ces scores n'ait réellement été validé, pouvaient amener à un faux sentiment de sécurité<sup>63,64</sup>. Sambecq et *al.* recommandaient la création d'un groupe de travail pour coordonner la mise en œuvre d'un SAP utilisable pour les hôpitaux généraux et universitaires<sup>63</sup>.

La conclusion de notre article était la suivante :

Cette étude prospective observationnelle multicentrique a montré que trois SAP (le PAWS, le PEWS et le Bedside PEWS), initialement développés pour les enfants admis en soins courants ou aux urgences pédiatriques pouvaient être utilisés dans les USC pédiatriques pour prédire le besoin d'une intervention médicale. D'autres études sont nécessaires pour valider leur utilisation courante avec une règle de prédiction d'appel du médecin en cas de détérioration dans ces USC pédiatriques.

# **Classification et caractérisation des patients d'USC pédiatriques selon l'évolution de la gravité de leur maladie**

Ce chapitre fait l'objet d'un article qui est en cours de rédaction en langue anglaise. L'article sera soumis à une revue internationale anglophone à comité de lecture.

De nombreuses USC pédiatriques ont été créées ces 10 dernières années. En 2007, une enquête téléphonique réalisée auprès des hôpitaux pédiatriques américains et canadiens révélait que 55% des hôpitaux étaient dotés d'une USC<sup>65</sup> mais la littérature décrivant ces unités et leur activité est restée pauvre. Une revue sur l'utilisation des USC a révélé une diversité de formats, mais les dénominateurs communs étaient la surveillance continue et l'assistance respiratoire, sans ventilation mécanique ni médicaments à action vasoactive<sup>62</sup>. L'objectif de ce travail était de décrire l'activité de ces USC et les patients hospitalisés dans ces unités et d'identifier les différentes trajectoires de gravité des enfants hospitalisés.

## **1- Méthodologie**

### **1.1- Schéma de l'étude et population**

Nous avons repris la cohorte prospective des sept USC pédiatriques d'hôpitaux du Nord-Pas de Calais, non attachées à une réanimation. Comme vu précédemment, les lits de ces USC sont équipés d'une station de surveillance centrale avec télémétrie continue. Elles sont également toutes équipées pour le traitement par oxygénothérapie à haut débit et 2 des 7 centres sont équipées pour la ventilation non invasive par pression positive continue nasale.

Tous les patients consécutifs âgés de moins de 18 ans admis dans les sept USC ont été inclus entre septembre 2012 et janvier 2014.

### **1.2- Données utilisées**

Les données utilisées pour ce deuxième objectif comprenaient l'âge, le sexe, la provenance de l'enfant (domicile, service d'urgence, service de pédiatrie, autres), les facteurs de comorbidité (affection chronique), la nature médicale ou chirurgicale de l'affection, la principale défaillance viscérale entraînant l'admission, la date et l'heure de sortie, la destination de l'enfant, le devenir à la sortie de l'USC (vivant/décédé).

Les paramètres de surveillance clinique comprenaient la température, le rythme cardiaque, la pression artérielle, le temps de recoloration cutanée, la saturation en oxygène, la fréquence respiratoire, le travail respiratoire (0 à 3), l'existence d'apnée, l'oxygénothérapie (air ambiant / canule / masque, débit et FIO2), le niveau de conscience (conscient / réponse à la voix / réponse à la douleur / inconscient). Ces paramètres étaient collectés par les infirmières sur des feuilles de surveillance journalière (standardisées pour l'étude dans les sept USC, Annexe 6) pour chaque patient à l'admission en USC (= H0) et au moins toutes les 8 heures (toutes les 8 heures de manière systématique et à chaque fois que l'infirmière suspectée une détérioration).

Des données relatives aux diagnostics codés à la fin du séjour en USC selon la Classification internationale des maladies 10<sup>ème</sup> édition et les actes codés selon la classification commune des actes médicaux (CCAM) étaient également utilisées pour ce deuxième objectif. Ces données étaient extraites à partir du Département d'Informatique Médical (DIM) de chaque hôpital participant. Une liste associant les noms des patients et les codes d'anonymisation individuels était fournie pour que les DIM puissent rendre un fichier anonyme (tableau A5 et A6, Annexe 14). Les DIM fournissaient également un fichier comprenant les données administratives relatives à chaque patient à des fins de vérification.

Les données relatives aux diagnostics et actes étaient listées à partir de l'Arrêté du 23 février 2013 relatif à la classification et à la prise en charge des prestations d'hospitalisation pour les activités de médecine, chirurgie, obstétrique et odontologie et pris en application de l'article L. 162-22-6 du code de la sécurité sociale.

Dans cet arrêté, les diagnostics et actes associés étaient explicités sous la forme de trois listes détaillées : \_\_\_\_\_

- Liste 1 : diagnostics et actes associés autorisant la facturation d'un supplément de surveillance continue
- Liste 2 : actes autorisant la facturation d'un supplément de surveillance continue
- Liste 3 : diagnostics et actes associés autorisant la facturation d'un supplément de surveillance continue pour les enfants de moins de 18 ans.

Un supplément dénommé « surveillance continue » était donnée pour chaque journée où le patient était pris en charge dans une unité de surveillance continue si l'une des conditions suivantes était remplie :

- un diagnostic ou un acte de la liste 1 avait été relevé
- un diagnostic ou un acte de la liste 2 établie avait été relevé

- un diagnostic ou un acte de la liste 3 établie avait été relevé.

Pour chaque patient, le recueil des diagnostics et actes était effectué à partir des listes ci-dessus existantes dans chaque centre. Une liste synthétique incluant les diagnostics et les actes les plus fréquemment recueillis a été donnée à tous les centres participants.

### **1.3- Critère de jugement**

La gravité de la maladie était évaluée par le SAP mesuré tout au long du séjour pour chaque patient (à l'admission et au minimum toutes 8 heures). Pour le traitement des items manquants nécessaire au calcul des SAP, il a été décidé de considérer également les items manquants comme normaux et le score a été calculé en attribuant à ces items manquants « 0 point », à condition que le nombre d'items manquants soit strictement inférieur à 3.

### **1.4- Choix du SAP**

Dans notre précédent travail, nous avons montré que les trois SAP (PAWS, PEWS et Bedside PEWS) mis au point initialement pour les enfants admis dans les services de pédiatrie ou se présentant aux urgences pédiatriques pouvaient être utilisés dans les USC pédiatriques pour prédire la nécessité d'une intervention médicale du fait d'une détérioration clinique<sup>66</sup>. La comparaison des différents items constituant les scores sont présentés en Annexe 12 et le détail des trois SAP en Annexe 1,2 et 3.

Pour évaluer la gravité des patients dans ce travail, nous avons choisi de sélectionner le score PAWS pour les raisons suivantes :

- le score PAWS était celui pour lequel le nombre de score complet était le plus élevé (86,1%)
- le score PAWS comprenait une évaluation des 3 principales défaillances : respiratoire, cardiovasculaire et neurologique (contrairement au Bedside PEWS qui ne comprenait pas d'évaluation neurologique)
- le score PAWS ne comprenait pas de mesure de pression artérielle (contrairement au Bedside PEWS), qui est un paramètre non surveillé en continu, occasionnant souvent le réveil de l'enfant et donc plus fréquemment manquant (9,5% des pressions artérielles nécessaires au calcul du Bedside PEWS étaient manquantes)
- et enfin, en comparaison au PEWS, nous avons préféré le score PAWS, qui nous a semblé plus facile à relever et à calculer au lit du malade car chaque item qui le compose était pris en

compte séparément, contrairement au score PEWS qui combinait plusieurs items regroupés en 3 défaillances (respiratoire, cardiovasculaire et neurologique).

### 1.5- Analyses statistiques

Les variables continues étaient exprimées par leurs valeurs médianes et interquartiles et les variables qualitatives par leurs effectifs et pourcentages. Les comparaisons entre groupes étaient effectuées par un test de khi 2 ( $\chi^2$ ) ou d'un test exact de Fisher en cas d'effectifs insuffisants. Pour les variables qualitatives. L'analyse des variables quantitatives était réalisée par un test-T de Student en cas de loi normale ou par un test U de Mann Whitney dans le cas contraire. La normalité des distributions étaient évaluées graphiquement et par le test de Shapiro Wilk. La gravité de la maladie était évaluée par le score PAWS, mesuré de l'admission (toutes les 8 heures pour chaque patient et en cas de détérioration) jusqu'au 6<sup>ème</sup> jour (6 jours correspondant au 95ème percentile de la durée du séjour).

Un modèle mixte linéaire à classes latentes (LCLMM) a été utilisé pour identifier différentes classes (k classes en général) en fonction des trajectoires des valeurs du PAWS au cours du séjour. Ces différentes classes, si elles existent, correspondent à différents profils de gravité. L'analyse a été ajustée sur le score PAWS à l'admission. Dans ce modèle, on suppose qu'il existe k classes (donc une variable qualitative latente à k modalités) correspondant à des groupes homogènes selon l'évolution d'un paramètre quantitatif.

Pour le modèle linéaire mixte, une évolution linéaire en fonction du temps a été choisie, avec un « intercept » et une pente aléatoire. Selon le critère d'information d'Akaike, (en anglais Akaike information criterion ou AIC, il n'était pas nécessaire de modéliser une évolution plus complexe, comme une évolution quadratique.

Le nombre de classes (k) n'était pas connu à l'avance et 4 modèles ont été testés avec k=1 à 4. Nous avons utilisé l'approche exposée dans la revue de Lennon et *al.*, publié en 2018, proposant une approche systématique et rationnelle pour construire, interpréter et choisir le meilleur modèle lorsqu'on utilise une méthodologie de classification par trajectoire<sup>67</sup>, nous avons utilisés différents critères pour sélectionner le meilleur modèle :

- les critères d'information bayésien (Bayes Information Criterion, BIC) : le modèle avec le BIC le plus bas est préféré,
- les probabilités postérieures d'affectation aux classes : dans un bon modèle, elles doivent être supérieures à 70% pour chaque classe,
- la pertinence clinique des trajectoires.

Les caractéristiques démographiques et cliniques étaient comparées entre les classes obtenues par un test du Chi-2 ou un test exact de Fisher pour les variables qualitatives et par une analyse de la variance (ANOVA) pour les variables quantitatives. Une analyse descriptive des principaux diagnostics étaient faites pour chaque classe obtenue.

Une valeur de  $p < 0,05$  était considérée comme significative. L'analyse statistique a été réalisée à l'aide de SAS version 9.4 (SAS Institute, Cary, NC, États-Unis).

## 2- Résultats

Parmi 2909 patients identifiés consécutivement en USC pédiatriques, 2868 patients ont été inclus pour l'analyse des données (flow chart présenté en figure 1).

Les principales caractéristiques générales de la population sont rappelées dans le tableau 4. L'âge médian était jeune (29 mois [5-103]) et 35% sont âgés de moins d'un an). La médiane de la durée du séjour était courte (1 jour [1-3]). Les raisons de l'admission en USC étaient médicales pour 95% et 33% des enfants avaient des comorbidités. Les patients étaient admis essentiellement depuis les urgences (70%) et transférés dans un service de soins courants de pédiatrie pour 61%.

**Tableau 4: caractéristiques générales de la population**

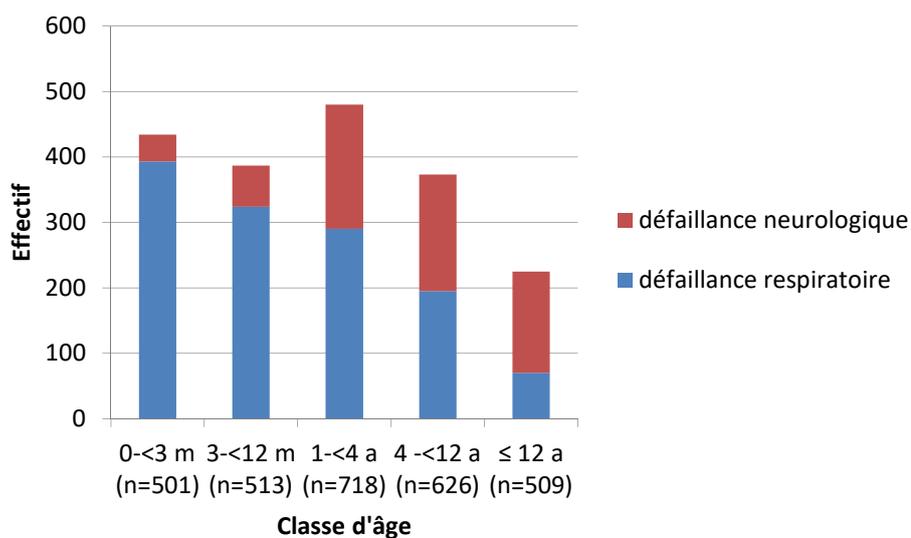
Caractéristiques	Total, n=2868
Sex ratio	1,26
Age (mois), mediane [IQ]	29 [5-103]
Age (mois), moyenne ( $\pm ET$ )	59 $\pm$ 66
Durée de séjour (jour), mediane [IQ]	1 [1-3]
Durée de séjour (jour), moyenne ( $\pm ET$ )	2 $\pm$ 2
Comorbidités, n (%)	945 (33)
Provenance:	
Urgences, n (%)	2006 (70)
Bloc, n (%)	71 (2)
Pédiatrie, n (%)	386 (13)
Reanimation pédiatrique, n (%)	25 (1)
Domicile, n (%)	377 (13)
Séjour médical, n (%)	2714 (95)
Chirurgical, n (%)	119 (4)
Destination :	
Pédiatrie, n (%)	1586 (55)
Domicile, n (%)	1024 (36)
Reanimation pédiatrique, n (%)	85 (3)
Décès	1

IQ : interquartile; ET: écart-type

La défaillance aiguë principale et l'étiologie motivant l'admission en USC pédiatrique sont présentées dans le tableau 5. Les principales défaillances étaient respiratoires (44%) et neurologiques (23%) et l'étiologie infectieuse était la plus fréquente représentant 52% des patients. La répartition de la défaillance principale par classe d'âge est présentée figure 2. La défaillance respiratoire était plus fréquente avant un an; la défaillance neurologique était plus fréquente après 12 ans et entre un an et 12 ans, la fréquence de la défaillance respiratoire diminuait avec l'âge (figure 2).

**Tableau 5: Défaillances aiguës principales et étiologies**

Défaillances aiguës et étiologies	Total, n=2868
Respiratoire, n (%)	1272 (44)
Infectieuse, n (%)	1007 (79)
Allergique, n (%)	89 (7)
Toxique, n (%)	25 (2)
Traumatique, n (%)	14 (1)
Autres, n (%)	137 (11)
Neurologique, n (%)	646 (23)
Infectieuse, n (%)	187 (29)
Traumatique, n (%)	70 (11)
Congénitale, n (%)	85 (13)
Toxique, n (%)	62 (10)
Autres, n (%)	242 (37)
Cardiovasculaire, n (%)	218 (8)
Infectieuse, n (%)	105 (48)
Traumatique, n (%)	12 (6)
Congénitale, n (%)	19 (9)
Toxique, n (%)	19 (9)
Autres, n (%)	63 (29)
Gastrointestinale, n (%)	183 (6)
Infectieuse, n (%)	116 (63)
Congénitale, n (%)	8 (4)
Toxique, n (%)	10 (5)
Autres, n (%)	49 (27)
Hématologique, n (%)	162 (6)
Oncologique, n (%)	50 (31)=
Infectieuse, n (%)	36 (22)
Immune, n (%)	12 (7)
Autres, n (%)	64 (40)
Endocrine, n (%)	124 (4)
Diabète, n (%)	98 (79)
Autres, n (%)	26 (21)
Hépatique, n (%)	39 (1)
Toxique, n (%)	28 (72)
Autres, n (%)	11 (28)
Genito-urinaire	30 (1)
Infectieuse, n (%)	18 (60)
Autres, n (%)	12 (40)
Polyviscérale	48 (2)
Infectieuse, n (%)	19 (40)
Autres, n (%)	29 (60)
Autres	146 (5)



m : mois, a : ans

**Figure 2 : Défaillance aiguë principale selon les différentes classes d'âge**

Les 10 principaux diagnostics et les 10 principaux actes codés à la fin du séjour en USC pédiatriques sont présentés tableau 6 et 7. Les 10 premiers diagnostics les plus fréquents représentaient plus de la moitié des patients (58%). La surveillance continue par scope était codée pour 90% des patients. L'oxygénothérapie était codée pour 20% des patients.

**Tableau 6: les 10 diagnostics les plus codés pour les séjours en USC pédiatrique**

DIAGNOSTICS		n=1662/2868 (58%)
1	Bronchiolite, n (%)	368 (12,8%)
2	Asthme, n (%)	285 (9,9%)
3	Insuffisance respiratoire aiguë, n (%)	198 (6,9%)
4	Etat de mal convulsif, n (%)	162 (5,6%)
5	Convulsions, n (%)	151 (5,3%)
6	Pneumopathie, n (%)	132 (4,6%)
7	Intoxication médicamenteuse, n (%)	123 (4,3%)
8	Diabète de type 1, n (%)	98 (3,4%)
9	Gastro-entérites et colites, n (%)	88 (3,1%)
10	Traumatisme crânien, n (%)	57 (2,0%)

**Tableau 7 : Les 10 actes les plus fréquemment codés pour les séjours en USC pédiatriques**

	Code Acte	Libellé	n= 8107
			n (%)
1	DEQP007	Surveillance continue de l'électrocardiogramme par oscilloscopie et/ou télésurveillance, avec surveillance continue de la pression intraartérielle et/ou de la saturation artérielle en oxygène par méthodes non effractives, par 24 heures	1583 (19,5%)
2	ZBQK002	Radiographie du thorax	1127 (13,9%)
3	DEQP004	Surveillance continue de l'électrocardiogramme par oscilloscopie et/ou télésurveillance, par 24 heures	986 (12,2%)
4	HSLF002	Alimentation parentérale avec apport de 20 à 35 kilocalories par kilogramme par jour [kcal/kg/jour], par 24 heures	777 (9,6%)
5	GLLD017	Oxygénothérapie avec surveillance continue de l'oxymétrie, en dehors de la ventilation mécanique, par 24 heures	573 (7,1%)
6	GELD005	Nébulisation d'agent thérapeutique à destination bronchique [aérosol] avec surveillance de la saturation en oxygène par mesure transcutanée [SpO2] et surveillance de la fréquence cardiaque, pendant au moins 2 heures	484 (6%)
7	GLQP001	Réglage du débit d'oxygène par surveillance transcutanée de la saturation en oxygène, pour instauration ou adaptation d'une oxygénothérapie de débit défini chez l'enfant par 24 heures	240 (3%)
8	GLHF001	Prélèvement de sang artériel avec gazométrie sanguine et mesure du pH, sans épreuve d'hyperoxie	186 (2,3%)
9	ACQK001	Scanographie du crâne et de son contenu, sans injection de produit de contraste	139 (1,7%)
10	DEQP003	Électrocardiographie sur au moins 12 dérivations	115 (1,4%)

Le tableau 8 présente les caractéristiques des patients transférés en réanimation pédiatrique depuis les services d'USC. Les patients transférés étaient significativement plus jeunes et leur séjour en USC plus court. La défaillance respiratoire et l'étiologie infectieuse étaient significativement plus fréquentes chez les patients transférés en réanimation pédiatrique.

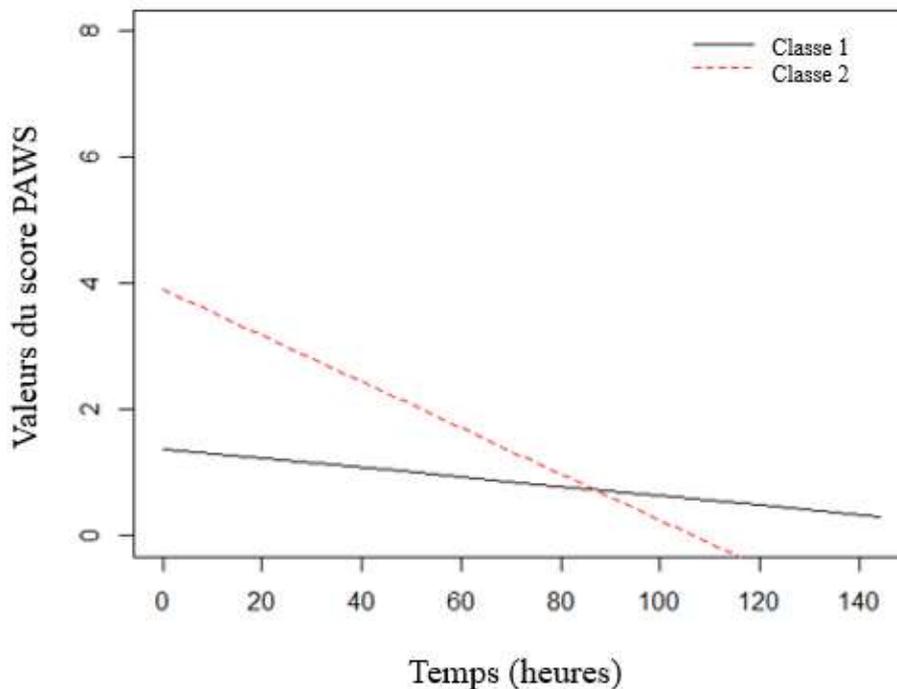
**Tableau 8 : Caractéristiques des patients transférés en réanimation pédiatrique**

Caractéristiques	Non transférés en réanimation n= 2783	Transférés en réanimation, n=85	p
Sex ratio	1,26	1,31	ns
Age (mois), médiane [IQ]	30[7-105]	5 [1-36]	<10 <sup>-4</sup>
Durée de séjour (jour), médiane [IQ]	1 [1-3]	1 [0-1]	<10 <sup>-4</sup>
Défaillance respiratoire, n (%)	1213 (44)	59 (69)	<10 <sup>-4</sup>
Infection, n (%)	1426 (51)	62 (73)	0,004
Comorbidités, n (%)	916 (33)	29 (34)	ns

ns: non significatif; IQ : interquartile

L'analyse en classe latente était réalisée sur 2636 patients car 234 patients n'avaient pas de score PAWS à l'admission (n=27) ou d'autres scores PAWS calculés en dehors de celui de l'admission (n=205) (Flow chart, figure 1).

Les figures de 3 à 5 représentent les trajectoires du score PAWS dans les différents modèles (à 2 classes, 3 classes et 4 classes).

**Figure 3 : Modèle à 2 classes de patients identifiés avec le PAWS**

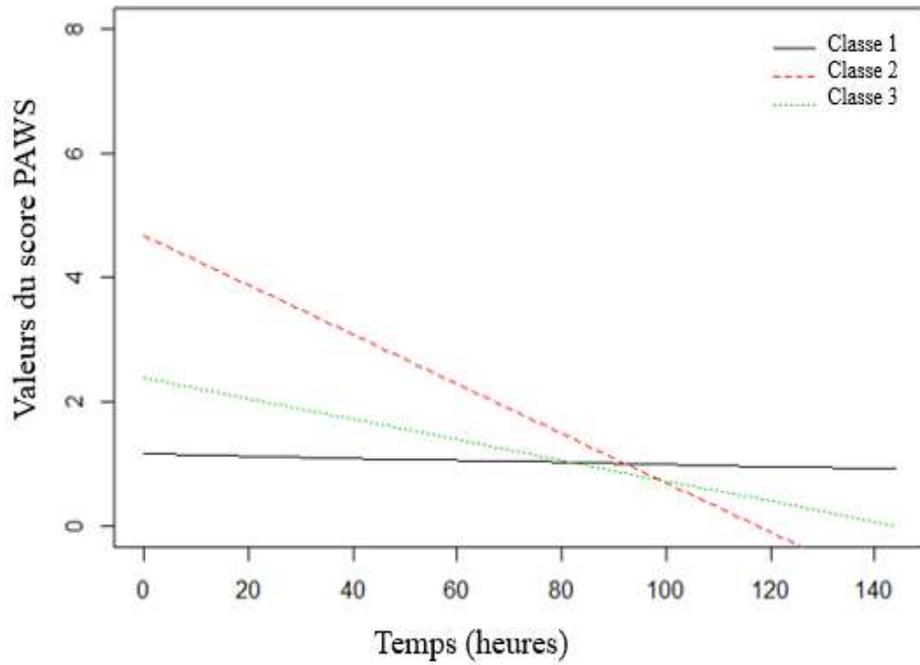


Figure 4 : Modèle à 3 classes de patients identifiés avec le PAWS

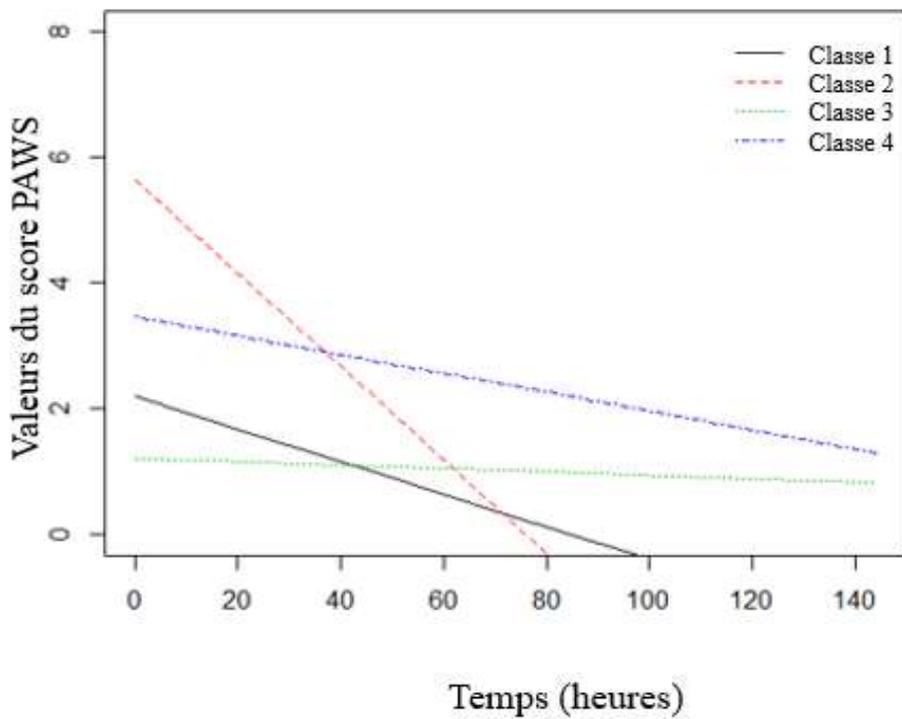


Figure 5: Modèle à 4 classes de patients identifiés avec le PAWS

Afin de déterminer le meilleur nombre de classes, nous avons utilisé les critères d'information bayésien. Les résultats sont présentés tableau 9.

**Tableau 9: Critères d'information bayésien en fonction du nombre de classe (K=1 à 4)**

Nombre de classe	Nombres de paramètres	BIC	Fréquence de la classe (%)			
			1	2	3	4
1 classe	7	52586	2636 (100)			
2 classes	11	52354	2296 (87,1)	340 (12,9)		
3 classes	15	52262	1593 (60,4)	170 (6,5)	873 (33,1)	
4 classes	19	52234	471 (17,9)	72 (2,7)	1736 (65,9)	357 (13,5)

BIC : bayesian information criteria

Le BIC le plus bas était pour le modèle à 4 classes : 52234 versus 52262 pour le modèle à 3 classes, ce qui est à peine plus élevé.

Concernant les probabilités postérieures à la classification (correspondant aux probabilités d'appartenance à la classe), pour le modèle à 3 classes, elles étaient de 85,7%, 82,4% et 76,8% soit toutes supérieures à 70%. Par contre pour le modèle à 4 classes, les probabilités postérieures à la classification n'étaient pas toutes supérieures à 70%, (67,0%, 75,5%, 86,2%, 70,2%).

Enfin, concernant la pertinence clinique, le modèle à 3 classes semblait plus pertinent puisque qu'il séparait les patients en 3 groupes de gravité :

3 groupes de patients :

- Classe 1 dit « stable »: 1593 patients peu graves à l'admission, qui restaient peu graves et stables sur le reste de l'hospitalisation
- Classe 2 dit « amélioration rapide » : 170 patients très graves à l'admission qui s'amélioraient rapidement
- Classe 3 dit « amélioration lente » : 873 patients moyennement graves qui s'amélioraient plus lentement

Dans le modèle à 4 classes, la classe 1 et la classe 4, avaient une trajectoire un peu similaire « Patients moyennement graves qui s'améliorent plus ou moins lentement ».

Le modèle retenu était donc le modèle à 3 classes.

Les caractéristiques générales associées aux différents profils de patients identifiés dans le modèle à 3 classes sont présentées tableau 10.

**Tableau 10: Comparaison des caractéristiques générales selon les différents profils de patients**

	Profil			p
	« stable » n=1593(60,4%)	« amélioration rapide » n=170 (6,5%)	« amélioration lente » n=873 (33,1%)	
Sexe ratio	1,2	1,3	1,4	0,060
Âge médian [IQ], mois	55 [15-144]	6 [2-16]	10 [4-36]	< 10 <sup>-4</sup>
Durée de séjour médiane [IQ], jours	1,1 [0,7-2,0]	2,4 [1,0-3,8]	1,9 [1,0-3,3]	< 10 <sup>-4</sup>
Comorbidités	566 (35%)	50 (29%)	260 (30%)	0,008
Provenance: urgences	1189 (75%)	91 (53%)	571 (66%)	<10 <sup>-4</sup>
Défaillance principale				< 10 <sup>-4</sup>
respiratoire	394 (25%)	150 (88%)	674 (74%)	
neurologique	466 (29%)	9 (5%)	105 (12%)	
cardio-circulatoire	165 (10%)	4 (2%)	29 (3%)	
gastro-intestinal	170 (11%)	2 (1%)	30 (3%)	
Étiologie infectieuse	623 (39%)	148 (87%)	679 (78%)	< 10 <sup>-4</sup>
Destination: pédiatrie (soins courants)	871 (55%)	81 (48%)	535 (62%)	<10 <sup>-3</sup>

IQ : interquartile

Les 3 groupes de patients étaient significativement différents. Le profil « stable » contenait le plus grand nombre de patient (60,4%). Les patients dans ce groupe étaient plus âgés (55 mois versus 6 mois pour le groupe « amélioration rapide » et 10 mois pour le groupe « amélioration lente »). La durée de séjour était plus courte pour le profil « stable » (1,1 jours versus 2,4 et 1,9 respectivement pour les 2 autres profils). La défaillance principale dans les groupes « amélioration rapide » ou « amélioration lente » était plus fréquemment respiratoire alors que la défaillance neurologique, cardio-circulatoire et gastro-intestinale était plus fréquente dans le profil « stable ». L'étiologie infectieuse était plus fréquente dans les groupes « amélioration rapide » ou « amélioration lente ».

Les 10 principaux diagnostics codés à la fin du séjour (selon la Classification internationale des maladies 10<sup>ème</sup> révision : CIM10) pour chaque profil de patients sont présentés tableau 11.

Dans le profil « stable », la variété de diagnostic était plus grande puisque les 10 diagnostics les plus fréquents ne représentaient que 44,1% des diagnostics principaux codés. Pour les deux autres profils, la variété de diagnostics présentés par les patients était bien plus faible puisque les 10 diagnostics les plus fréquents représentaient 85,3% pour le profil « amélioration rapide » et 69,2% pour le profil « amélioration lente ». Pour le groupe « stable », les trois premiers diagnostics étaient l'asthme, l'intoxication médicamenteuse et les convulsions et représentaient à eux trois 18,4% des diagnostics codés. Pour les profils « amélioration rapide » et « amélioration lente », les trois premiers diagnostics étaient les mêmes pour les deux profils : la bronchiolite, l'insuffisance respiratoire aiguë et l'asthme et représentaient 69,4% et 48,8% des diagnostics codés respectivement. Pour ces deux derniers profils, la défaillance concernée par les quatre diagnostics les plus fréquents étaient respiratoire.

**Tableau 11 : Principaux diagnostics dans chaque profil de patients**

Profil	Diagnostics	nombre (%)
« stable »		(n=1593)
1	Asthme	100 (6,3)
2	Intoxication médicamenteuse	99 (6,2)
3	Convulsions	94 (5,9)
4	Diabète insulino-dépendant	84 (5,3)
5	Etat de mal convulsif	77 (4,8)
6	Bronchiolite	61 (3,8)
7	Gastroentérite	60 (3,8)
8	Insuffisance respiratoire aiguë	47 (2,9)
9	Traumatisme crânien	40 (2,5)
10	Pneumopathie	40 (2,5)
« amélioration rapide »		(n=170)
1	Bronchiolite	63 (37,1)
2	Insuffisance respiratoire aiguë	31 (18,2)
3	Asthme	24 (14,1)
4	Pneumopathie	11 (6,5)
5	Etat de mal convulsif	4 (2,3)
6	Convulsions	3 (1,8)
7	Coqueluche	3 (1,8)
8	Diabète insulino-dépendant	2 (1,2)
9	Atélectasie	2 (1,2)
10	Allergie alimentaire	2 (1,2)
« amélioration lente »		(n=873)
1	Bronchiolite	219 (25,1)
2	Insuffisance respiratoire aiguë	111 (12,7)
3	Asthme	96 (11,0)
4	Pneumopathie	72 (8,2)
5	Convulsions	42 (4,8)
6	Rhinopharyngite	16 (1,8)
7	Laryngite	15 (1,7)
8	Etat de mal convulsif	13 (1,5)
9	Gastroentérite	11 (1,3)
10	Diabète insulino-dépendant	9 (1,0)

### 3- Discussion

Dans cette cohorte de 2868 enfants admis dans sept USC d'hôpitaux généraux, les patients étaient jeunes (médiane à 29 mois) et la durée du séjour était courte (médiane à 1 jour). Les deux principales défaillances à l'origine de l'hospitalisation en USC étaient respiratoire (44%) et neurologique (22%). L'étiologie principale était infectieuse (52%). Le taux de transfert en réanimation était de 3%. Il existait 3 profils de patient : le profil « stable » (60,4%), les profils « amélioration rapide » (6,5%) et « amélioration lente » (33,1%). Les patients du profil « stable » étaient plus âgés et avaient une durée de séjour plus courte et avaient une défaillance neurologique principalement. Dans le profil « stable », les diagnostics étaient beaucoup plus variés. Les principaux diagnostics des patients « amélioration rapide » ou « amélioration lente » étaient la bronchiolite et l'asthme.

Il existe une grande hétérogénéité dans les caractéristiques des USC. Celles-ci peuvent être chirurgicale, médicale, mixte. Certaines sont intégrées au sein d'un service de réanimation ou encore adjacentes au service de réanimation, d'autres sont indépendantes, ou font parties d'un service de soins courants. Comme rapporté par Plate et *al.*, les points communs entre ces unités sont le monitoring cardio-respiratoire et l'assistance respiratoire non invasive<sup>62</sup>. Les unités attachées à un service de réanimation semblent moins limitées que les unités séparées, notamment pour l'administration d'amine vasoactive ou la ventilation non invasive. Les caractéristiques des patients et la gravité des patients dans les USC pédiatriques attachées à une réanimation sont donc probablement assez différentes des patients des USC séparées d'une réanimation.

Il existe peu d'études indiquant les caractéristiques des enfants hospitalisés en USC pédiatriques que ce soit attaché à une réanimation<sup>68,69</sup> ou séparée<sup>70,71</sup>.

#### 3.1- Caractéristiques générales, défaillances aiguës et étiologies

Dans la cohorte de Hubert et *al.*<sup>68</sup>, incluant 460 enfants admis dans une USC attachée à une réanimation dans un hôpital universitaire, l'âge moyen était de 61,8 mois ( $\pm$  63,3) et la durée moyenne du séjour de 3,6 jours ( $\pm$  6). Quarante-sept pour cent des patients avaient des comorbidités. Le type de séjour était chirurgical dans 47,2% des cas. Les défaillances principales étaient essentiellement neurologiques pour 35,9% et respiratoires pour 25,6%.

Dans la cohorte Hamze-sinno et *al.*<sup>69</sup>, comprenant 912 enfants également admis dans une USC attachée à une réanimation dans un hôpital universitaire, l'âge médian était de 24,5 [4,0 à 91,8] et la durée de séjour médiane de 2 jours [2-3]. Les défaillances principales étaient respiratoires

pour 36,9%, neurologiques pour 16,4% et cardiovasculaires pour 14,8%. Les patients étaient transférés en réanimation dans 9,6% des cas et 1,6% décédés.

Dans notre cohorte, l'âge était similaire (âge médian 29 mois et âge moyen 59 mois), mais la durée du séjour plus courte (médiane à 1 jour, moyenne à 2 jours). Nous avons peu de séjours chirurgicaux (5%), comparés au 47% dans la cohorte de Hubert et *al.*, ce qui s'explique par l'absence de chirurgie pédiatrique dans beaucoup d'hôpitaux généraux contrairement aux centres universitaires<sup>68</sup>.

Les principales défaillances étaient respiratoires et neurologiques dans notre étude, une prédominance respiratoire jusqu'à 12 ans qui tend à diminuer avec l'âge. L'insuffisance respiratoire est la principale défaillance dans la plupart des autres cohortes suivies de la défaillance neurologique<sup>69-71</sup>. Thiriez et *al.*, dans une enquête dans trois régions françaises estimant les besoins en USC pédiatriques rapportaient que 48% des affections responsables d'une demande de transfert en USC pédiatrique étaient dues à une affection respiratoire, suivies des affections neurologiques pour 17,9%<sup>72</sup>.

Dans la cohorte de Hubert et *al.*, la défaillance principale la plus fréquente était neurologique. Cette prédominance neurologique peut être expliquée par leur recrutement élevé de séjours chirurgicaux et leur taux élevé de comorbidité (47%)<sup>68</sup>.

Dans notre étude, l'étiologie prédominante était infectieuse pour 52% des enfants. La période d'inclusion s'étendait de septembre 2012 à janvier 2014 et comprenait plus de mois d'automne et d'hiver pouvant participer à cette prédominance infectieuse. La faible proportion de séjours chirurgicaux (5%) peut également expliquer la prédominance d'étiologie infectieuse.

### **3.2- Les patients transférés en réanimation**

Le taux de transfert en réanimation était plus bas dans notre étude (3%) comparé à la cohorte de Hamze-Sinno et *al.* (9,6%)<sup>69</sup> mais les patients hospitalisés dans les USC attachées à une réanimation sont probablement plus grave. Ces USC ont, comme le montrait Plate et *al.*, moins de limites concernant les thérapeutiques de suppléance comme l'administration d'amines ou la ventilation non invasive<sup>62</sup>. Dans la cohorte de Gatti et *al.*, le taux de patients transférés dans une réanimation était de 8%, mais cette étude était faite en Polynésie<sup>70</sup>. L'organisation des soins différente due à l'isolement géographique, le besoin de soins spécialisés a dû impacter le taux de transfert en réanimation.

Les patients transférés en réanimation depuis les USC, dans notre étude, étaient plus jeunes que les patients non transférés, comme observé dans la cohorte de hamze-sinno et *al.*<sup>69</sup>. La défaillance respiratoire et l'étiologie infectieuse était plus fréquente chez les patients transférés

dans notre cohorte. Hamze-sinno et *al.* objectivait également plus fréquemment la défaillance respiratoire chez les patients transférés<sup>69</sup>.

### 3.3- L'analyse en classe latente

L'analyse en classe latente avec l'utilisation d'un SAP, le PAWS, pour refléter la gravité des patients, a permis dans notre étude de mettre en évidence 3 profils de patients en USC pédiatriques : le profil « stable », les profils « amélioration rapide » et « amélioration lente ». Les patients de ces deux derniers groupes avaient un PAWS plus élevé en début de séjour tendant à s'améliorer plus ou moins rapidement contrairement au groupe « stable » qui avait un PAWS bas et qui le restait. Le profil « stable » avait également une durée de séjour plus courte. Les caractéristiques des patients des 2 profils « amélioration » étaient les mêmes que pour les patients transférés en réanimation, soit des patients plus jeunes, présentant une défaillance respiratoire et une étiologie infectieuse. Trois diagnostics prédominaient dans ces deux groupes : la bronchiolite, l'insuffisance respiratoire et l'asthme.

La bronchiolite, l'asthme, les convulsions, les intoxications médicamenteuses, la découverte de diabète insulino-dépendant sont des diagnostics pour lesquels l'USC est probablement le plus utile. En effet, ces maladies sont aiguës, peuvent être sévères, et sont à haut risque de détérioration avec un faible risque de mortalité<sup>73</sup>. Russ et *al.*, par le biais de six scénarios de patients, envoyés à 192 hôpitaux, voulaient évaluer si le triage de patients nécessitant une surveillance continue sans recours à des traitements de suppléance invasif, donc candidats à une orientation vers l'USC, était correct. Dans cette étude, trois des six scénarios de patients proposés aux médecins, étaient la bronchiolite, l'asthme, et le diabète insulino-dépendant<sup>74</sup>. Ces trois diagnostics représentaient 1/3 des diagnostics des patients d'USC dans notre étude.

Ces USC sont également très utiles pour les enfants jeunes, plutôt atteints de défaillance respiratoire et à plus haut à risque de détérioration, nécessitant plus fréquemment un transfert en réanimation dans notre étude. Thiriez et *al.*, dans son enquête concernant toutes les services pédiatriques de trois régions de France afin d'estimer les besoins en USC pédiatriques et les caractéristiques des patients qui auraient pu relever de celles-ci, rapportaient que 1/3 des enfants avait moins de 1 an et 1/4 moins de 6 mois<sup>72</sup>. L'essentiel des affections responsables d'une demande de transfert en USC de la part des services de pédiatrie médicale était respiratoire (47,4 %), avec surtout des bronchiolites, des pneumopathies et des crises d'asthme<sup>72</sup>. Notre étude confirme les estimations des besoins en USC pédiatriques qui avaient été faites dans cette enquête.

### 3.4- Limites

Une des limites de ce travail est l'exhaustivité et l'exactitude des diagnostics et actes codés qui ne peuvent être vérifiés. Il n'existe aucune formation au codage durant les études médicales. Cette tâche revient dans la plupart des centres aux médecins. Certaines structures hospitalières font parfois appel à des sociétés privées spécialisées. Ce n'était pas le cas dans notre étude pour aucun des centres participants. De même, le diagnostic « insuffisance respiratoire aiguë » souvent codé dans notre étude (6,9%) est peu précis et peut avoir sous-estimé la proportion de bronchiolite, d'asthme, de pneumopathie ou de toute autre maladie occasionnant une insuffisance respiratoire aiguë.

Une autre limite est l'exclusion des patients n'ayant pas de score mesuré après H0. Ces patients n'avaient pas de scores à H8, impliquant un séjour très court. Certains de ces patients peuvent avoir été transférés rapidement en réanimation pédiatrique du fait de leur gravité. L'exclusion de ces patients pourrait expliquer l'absence de profil « aggravation » identifié dans notre modèle.

Enfin, cette étude a été réalisée au sein de sept USC non attachées à une réanimation. Les caractéristiques et les différents profils de patients ne sont pas généralisables aux USC attachées aux réanimations. Les patients de ces USC attachées aux réanimations pourraient être plus graves, présenter des pathologies différentes du fait de spécialités (chirurgicales, oncologiques, de greffe, centres de référence de maladies rares...) présentes dans ces centres, souvent universitaires, dotés d'une réanimation pédiatrique.

## Conclusion

Cette étude observationnelle multicentrique prospective a permis de valider l'utilisation des SAP, à travers trois SAP (PAWS, PEWS et Bedside PEWS) mis au point initialement pour les enfants admis dans les services de pédiatrie ou se présentant aux urgences pédiatriques, dans les USC pédiatriques pour prédire la nécessité d'une intervention médicale. Nous proposons une approche différente grâce à une méthodologie prenant en compte les mesures répétées chez un même patient permettant d'utiliser le SAP avec une règle de prédiction d'appel du médecin par l'infirmière en cas de détérioration clinique basée sur le score à l'admission, le score au temps t et le temps t. Des études complémentaires seront nécessaires dans différentes USC et dans différents pays pour valider l'utilisation des SAP avec une règle de prédiction d'appel du médecin pour détérioration clinique.

Un des SAP, le PAWS, validé précédemment, a été utilisé pour classifier les malades d'USC selon différentes trajectoires de gravité. L'analyse en classe latente mettait en évidence différentes trajectoires de gravité : le profil « stable » (60,4%), le profil « amélioration rapide » (6,5%) et le profil « amélioration lente » (33,1%). Bien que ces USC aient été développées pour des patients nécessitant une surveillance continue du fait d'un haut risque de dégradation, on constate que la majorité des patients hospitalisés dans ces unités appartiennent au profil « stable ». Nous avons pu montrer que seulement 10 diagnostics représentaient plus de la moitié des patients. La défaillance principale à l'origine de l'admission en USC était respiratoire (44%) et l'étiologie infectieuse était la plus fréquente (52%). Les deux diagnostics les plus fréquents étaient l'asthme et la bronchiolite. Ce travail confirme également que ces USC pédiatriques sont très utiles pour les enfants jeunes, atteint d'une défaillance respiratoire car ces patients sont les patients les plus sévères (profil « amélioration rapide » et « amélioration lente ») et les plus à risque d'une détérioration nécessitant d'un transfert en réanimation. Enfin, la durée de séjour en USC pédiatriques courte (1 jour [1-3]), soutient bien le fait que ces unités sont « intermédiaires » et l'orientation vers un niveau de soins plus élevé en cas de détérioration ou un service de soins courants en cas de stabilisation se fait rapidement.

## Références

1. Dager S, Michot C, Pinto Da Costa N. Continuous monitoring unit in a university hospital: an operating survey. *Arch Pediatr* 2008;15:683-685.
2. Jaimovich DG, American Academy of Pediatrics Committee on Hospital Care and Section on Critical Care. Admission and discharge guidelines for the pediatric patient requiring intermediate care. *Pediatrics*. 2004;113:1430-1433.
3. Leclerc F, Leteurtre S, Binoche A. Pediatric intermediate care: official French texts. *Arch Pediatr*. 2008;15:679-682.
4. Décret N° 2006-72 Du 24 Janvier 2006 Relatif à La Réanimation Dans Les Établissements de Santé et Modifiant Le Code de La Santé Publique (Dispositions Réglementaires); 2006.
5. Shann F, Pearson G, Slater A, Wilkinson K. Paediatric index of mortality (PIM): a mortality prediction model for children in intensive care. *Intensive Care Med*. 1997;23:201-207.
6. Yeh TS, Pollack MM, Ruttimann UE, Holbrook PR, Fields AI. Validation of a physiologic stability index for use in critically ill infants and children. *Pediatr Res*. 1984;18:445-451.
7. Leteurtre S, Duhamel A, Grandbastien B, et al. Daily estimation of the severity of multiple organ dysfunction syndrome in critically ill children. *CMAJ Can Med Assoc J J Assoc Medicale Can*. 2010;182:1181-1187.
8. Leteurtre S, Lampin M-E, Grandbastien B, Recher M, Duhamel A. Les scores de gravité généraux et de dysfonctions d'organes en réanimation pédiatrique : quoi de neuf en 2016 ? *Médecine Intensive Réanimation*. 2016;25:604-618.
9. Gulla KM, Sachdev A. Illness severity and organ dysfunction scoring in Pediatric Intensive Care Unit. *Indian J Crit Care Med Peer-Rev Off Publ Indian Soc Crit Care Med*. 2016;20:27-35.
10. Lemeshow S, Le Gall JR. Modeling the severity of illness of ICU patients. A systems update. *JAMA*. 1994;272:1049-1055.
11. Royston P, Sauerbrei W, Altman DG. Modeling the effects of continuous risk factors. *J Clin Epidemiol*. 2000;53:219-221.
12. Shann F. Are we doing a good job: PRISM, PIM and all that. *Intensive Care Med*. 2002;28:105-107.
13. Keegan MT, Gajic O, Afessa B. Severity of illness scoring systems in the intensive care unit. *Crit Care Med*. 2011;39:163-169.
14. Lemeshow S, Hosmer DW. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol*. 1982;115:92-106.

15. Chapman SM, Grocott MPW, Franck LS. Systematic review of paediatric alert criteria for identifying hospitalised children at risk of critical deterioration. *Intensive Care Med.* 2010;36:600-611.
16. Berg MD, Nadkarni VM, Zuercher M, Berg RA. In-hospital pediatric cardiac arrest. *Pediatr Clin North Am.* 2008;55:589-604.
17. Lambert V, Matthews A, MacDonell R, Fitzsimons J. Paediatric early warning systems for detecting and responding to clinical deterioration in children: a systematic review. *BMJ Open.* 2017;7:e014497.
18. Olson D, Davis NL, Milazi R, et al. Development of a severity of illness scoring system (inpatient triage, assessment and treatment) for resource-constrained hospitals in developing countries. *Trop Med Int Health TM IH.* 2013;18:871-878.
19. Akre M, Finkelstein M, Erickson M, Liu M, Vanderbilt L, Billman G. Sensitivity of the pediatric early warning score to identify patient deterioration. *Pediatrics.* 2010;125:e763-769.
20. Robson M-AJ, Cooper CL, Medicus LA, Quintero MJ, Zuniga SA. Comparison of three acute care pediatric early warning scoring tools. *J Pediatr Nurs.* 2013;28:e33-41.
21. Duncan H, Hutchison J, Parshuram CS. The Pediatric Early Warning System score: a severity of illness score to predict urgent medical need in hospitalized children. *J Crit Care.* 2006;21:271-278.
22. Tucker KM, Brewer TL, Baker RB, Demeritt B, Vossmeier MT. Prospective evaluation of a pediatric inpatient early warning scoring system. *J Spec Pediatr Nurs JSPN.* 2009;14:79-85.
23. Parshuram CS, Hutchison J, Middaugh K. Development and initial validation of the Bedside Paediatric Early Warning System score. *Crit Care Lond Engl.* 2009;13:R135.
24. Parshuram CS, Duncan HP, Joffe AR, et al. Multicentre validation of the bedside paediatric early warning system score: a severity of illness score to detect evolving critical illness in hospitalised children. *Crit Care Lond Engl.* 2011;15:R184.
25. Fuijkschot J, Vernhout B, Lemson J, Draaisma JMT, Loeffen JLCM. Validation of a Paediatric Early Warning Score: first results and implications of usage. *Eur J Pediatr.* 2015;174:15-21.
26. Mandell IM, Bynum F, Marshall L, Bart R, Gold JI, Rubin S. Pediatric Early Warning Score and unplanned readmission to the pediatric intensive care unit. *J Crit Care.* 2015;30:1090-1095.
27. Skaletzky SM, Raszynski A, Totapally BR. Validation of a modified pediatric early warning system score: a retrospective case-control study. *Clin Pediatr (Phila).* 2012;51:431-435.
28. Bell D, Mac A, Ochoa Y, Gordon M, Gregurich MA, Taylor T. The Texas Children's Hospital Pediatric Advanced Warning Score as a predictor of clinical deterioration in

- hospitalized infants and children: a modification of the PEWS tool. *J Pediatr Nurs*. 2013;28:e2-9.
29. Parshuram CS, Bayliss A, Reimer J, Middaugh K, Blanchard N. Implementing the Bedside Paediatric Early Warning System in a community hospital: A prospective observational study. *Paediatr Child Health*. 2011;16:e18-22.
  30. McKay H, Mitchell IA, Sinn K, Mugridge H, Lafferty T, Van Leuvan C, Mamootil S, Abdel-Latif ME. Effect of a multifaceted intervention on documentation of vital signs and staff communication regarding deteriorating paediatric patients. *J Paediatr Child Health*. 2013;49:48-56.
  31. Randhawa S, Roberts-Turner R, Woronick K, DuVal J. Implementing and sustaining evidence-based nursing practice to reduce pediatric cardiopulmonary arrest. *West J Nurs Res*. 2011;33:443-456.
  32. Gerry S, Birks J, Bonnici T, Watkinson PJ, Kirtley S, Collins GS. Early warning scores for detecting deterioration in adult hospital patients: a systematic review protocol. *BMJ Open*. 2017;7:e019268.
  33. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. 2015;162:55-63.
  34. Chapman SM, Wray J, Oulton K, Peters MJ. Systematic review of paediatric track and trigger systems for hospitalised children. *Resuscitation*. 2016;109:87-109.
  35. Auriant I, Vinatier I, Thaler F, Tourneur M, Loirat P. Simplified acute physiology score II for measuring severity of illness in intermediate care units. *Crit Care Med*. 1998;26:1368-1371.
  36. Prin M, Harrison D, Rowan K, Wunsch H. Epidemiology of admissions to 11 stand-alone high-dependency care units in the UK. *Intensive Care Med*. 2015;41:1903-1910.
  37. Lucena JF, Alegre F, Rodil R, Landecho MF, García-Mouriz A, Marqués M, Aquerreta I, García N, Quiroga J. Results of a retrospective observational study of intermediate care staffed by hospitalists: impact on mortality, co-management, and teaching. *J Hosp Med*. 2012;7:411-415.
  38. Hager DN, Tanykonda V, Noorain Z, Sahetya SK, Simpson CE, Lucena JF, Needham DM. Hospital mortality prediction for intermediate care patients: Assessing the generalizability of the Intermediate Care Unit Severity Score (IMCUSS). *J Crit Care*. 2018;46:94-98.
  39. Alegre F, Landecho MF, Huerta A, Fernández-Ros N, Martínez-Urbistondo D, García N, Quiroga J, Lucena JF. Design and Performance of a New Severity Score for Intermediate Care. *PLoS One*. 2015;10:e0130989.
  40. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA*. 1993;270:2957-2963.

41. Moreno RP, Metnitz PGH, Almeida E, et al. SAPS 3--From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission. *Intensive Care Med.* 2005;31:1345-1355.
42. Lucena JF, Alegre F, Martinez-Urbistondo D, Landecho MF, Huerta A, García-Mouriz A, García N, Quiroga J. Performance of SAPS II and SAPS 3 in intermediate care. *PloS One.* 2013;8:e77229.
43. Egdell P, Finlay L, Pedley DK. The PAWS score: validation of an early warning scoring system for the initial assessment of children in the emergency department. *Emerg Med J EMJ.* 2008;25:745-749.
44. Solevåg AL, Eggen EH, Schröder J, Nakstad B. Use of a modified pediatric early warning score in a department of pediatric and adolescent medicine. *PloS One.* 2013;8:e72534.
45. Bradman K, Borland M, Pascoe E. Predicting patient disposition in a paediatric emergency department. *J Paediatr Child Health.* 2014;50:E39-44.
46. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143:29-36.
47. Foulkes AS, Azzoni L, Li X, Johnson MA, Smith C, Mounzer K, Montaner LJ.. Prediction based classification for longitudinal biomarkers. *Ann Appl Stat.* 2010;4:1476-1497.
48. Trubey R, Huang C, Lugg-Widger FV, Hood K, Allen D, Edwards D, Lacy D, Lloyd A, Mann M, Mason B, Oliver A, Roland D, Sefton G, Skone R, Thomas-Jones E, Tume LN, Powell C. Validity and effectiveness of paediatric early warning systems and track and trigger tools for identifying and reducing clinical deterioration in hospitalised children: a systematic review. *BMJ Open.* 2019;9:e022105.
49. Agulnik A, Forbes PW, Stenquist N, Rodriguez-Galindo C, Kleinman M. Validation of a Pediatric Early Warning Score in Hospitalized Pediatric Oncology and Hematopoietic Stem Cell Transplant Patients. *Pediatr Crit Care Med J Soc Crit Care Med World Fed Pediatr Intensive Crit Care Soc.* 2016;17:e146-153.
50. Tume L. The deterioration of children in ward areas in a specialist children's hospital. *Nurs Crit Care.* 2007;12:12-19.
51. Bonafide CP, Roberts KE, Weirich CM, Paciotti B, Tibbetts KM, Keren R, Barg FK, Holmes JH. Beyond statistical prediction: qualitative evaluation of the mechanisms by which pediatric early warning scores impact patient safety. *J Hosp Med.* 2013;8:248-253.
52. Jensen CS, Nielsen PB, Olesen HV, Kirkegaard H, Aagaard H. Pediatric Early Warning Score Systems, Nurses Perspective - A Focus Group Study. *J Pediatr Nurs.* February 2018.
53. Gawronski O, Ciofi Degli Atti ML, Di Ciommo V, Cecchetti C, Bertaina A, Tiozzo E, Raponi M; Stem Cell Transplant Unit BedsidePEWS Study Group. Accuracy of Bedside

- Paediatric Early Warning System (BedsidePEWS) in a Pediatric Stem Cell Transplant Unit. *J Assoc Pediatr Oncol Nurses*. 2016;33:249-256.
54. McLellan MC, Gauvreau K, Connor JA. Validation of the Cardiac Children's Hospital Early Warning Score: an early warning scoring tool to prevent cardiopulmonary arrests in children with heart disease. *Congenit Heart Dis*. 2014;9:194-202.
  55. Gold DL, Mihalov LK, Cohen DM. Evaluating the Pediatric Early Warning Score (PEWS) system for admitted patients in the pediatric emergency department. *J Soc Acad Emerg Med*. 2014;21:1249-1256.
  56. Chaiyakulsil C, Pandee U. Validation of pediatric early warning score in pediatric emergency department. *Jpn Pediatr Soc*. 2015;57:694-698.
  57. Seiger N, Maconochie I, Oostenbrink R, Moll HA. Validity of different pediatric early warning scores in the emergency department. *Pediatrics*. 2013;132:e841-850.
  58. Chapman SM, Wray J, Oulton K, Peters MJ. "Death is not the answer": the challenge of measuring the impact of early warning systems. *Arch Dis Child*. 2019;104:210-211.
  59. Jensen CS, Olesen HV, Aagaard H, Svendsen MLO, Kirkegaard H. Comparison of Two Pediatric Early Warning Systems: A Randomized Trial. *J Pediatr Nurs*. 2019;44:e58-e65.
  60. Parshuram CS, Dryden-Palmer K, Farrell C, Gottesman R, Gray M, Hutchison JS, Helfaer M, Hunt EA, Joffe AR, Lacroix J, Moga MA, Nadkarni V, Ninis N, Parkin PC, Wensley D, Willan AR, Tomlinson GA; Canadian Critical Care Trials Group and the **EPOCH** Investigators. Effect of a Pediatric Early Warning System on All-Cause Mortality in Hospitalized Pediatric Patients: The EPOCH Randomized Clinical Trial. *JAMA*. 2018;319:1002-1012.
  61. Mulherin SA, Miller WC. Spectrum bias or spectrum effect? Subgroup variation in diagnostic test evaluation. *Ann Intern Med*. 2002;137:598-602.
  62. Plate JDJ, Leenen, Luke. P. H., Houwert, Marijn., Hietbrink, Falco. Utilisation of Intermediate Care Units: A Systematic Review. *Crit Care Res Pract*. 2017:1-10.
  63. Sambeek SJ van, Fuijkschot J, Kramer BW, Vos GD. Pediatric Early Warning System Scores: Lessons to be Learned. *J Pediatr Intensive Care*. 2018;7:27-32.
  64. Chapman SM, Maconochie IK. Early warning scores in paediatrics: an overview. *Arch Dis Child*. 2019;104:395-399.
  65. VandenBerg SD, Hutchison JS, Parshuram CS, Paediatric Early Warning System Investigators. A cross-sectional survey of levels of care and response mechanisms for evolving critical illness in hospitalized children. *Pediatrics*. 2007;119:e940-946.
  66. Lampin ME, Duhamel A, Behal H, Recher M, Leclerc F, Leteurtre S. Use of paediatric early warning scores in intermediate care units. *Arch Dis Child*. August 2019. doi:10.1136/archdischild-2019-317055

67. Lennon H, Kelly S, Sperrin M, et al. Framework to construct and interpret latent class trajectory modelling. *BMJ Open*. 2018;8:e020683.
68. Hubert A, Ford-Chessel C, Berthiller J, Peretti N, Javouhey E, Valla FV. [Nutritional status in pediatric intermediate care: Assessment at admission, progression during the stay and after discharge]. *Arch Pediatr* 2016;23:333-339.
69. Hamze-Sinno R, Abdoul H, Neve M, Tsapis M, Jones P, Dauter S. Can we easily anticipate on admission pediatric patient transfers from intermediate to intensive care? *Minerva Anesthesiol*. 2011;77:1022-1023.
70. Gatti H, Dauter S, Sommet J, Chenel C, Naudin J. Pediatric intermediate care unit in general hospital: recent survey in French Polynesia. *Arch Pediatr Organe Off Soc Francaise Pediatr*. 2014;21:272-278. doi:10.1016/j.arcped.2013.12.020
71. Hernandez X, Oyharcabal V, Jouvencel P. Place of paediatric continuous care unit in general hospital. *Arch Pediatr Organe Off Soc Francaise Pediatr*. 2008;15:686-688.
72. Thiriez G, Lefebvre A. Pediatric intermediate care demand, results of a survey in 3 French regions. *Arch Pediatr Organe Off Soc Francaise Pediatr*. 2010;17:1147-1152.
73. Slater A, Shann F, Pearson G, Paediatric Index of Mortality (PIM) Study Group. PIM2: a revised version of the Paediatric Index of Mortality. *Intensive Care Med*. 2003;29:278-285.
74. Russ CM, Agus M. Triage of Intermediate-Care Patients in Pediatric Hospitals. *Hosp Pediatr*. 2015;5:542-547.
75. Thomas-Jones E, Lloyd A, Roland D, Sefton G, Tume L, Hood K, Huang C, Edwards D, Oliver A, Skone R, Lacy D, Sinha I, Preston J, Mason B, Jacob N, Trubey R, Strange H, Moriarty Y, Grant A, Allen D, Powell C. A prospective, mixed-methods, before and after study to identify the evidence base for the core components of an effective Paediatric Early Warning System and the development of an implementation package containing those core recommendations for use in the UK: Paediatric early warning system - utilisation and mortality avoidance- the PUMA study protocol. *BMC Pediatr*. 2018;18:244.
76. High dependency care for children - time to move on (2014). RCPCH. <https://www.rcpch.ac.uk/resources/high-dependency-care-children-time-move-2014>. Accessed July 11, 2019.

## Annexes

### Annexe 1: The Paediatric advanced warning score (PAWS), Egdell et al

Respiratory rate	+30	3	age (years)	Normal range for respiratory rate		
	+20	2			<1	30-40
	+10	1			1-2	25-35
	normale	0			2-5	25-30
	-10	2			5-12	20-25
			>12	15-20		
Work of breathing	Severe increase	3				
	Moderate increase	2				
	normal	0				
Oxygen saturation (%)	>ou= 93%	0				
	90-92	1				
	85-89	2				
	<85	3				
Temprature (°C)	39	2				
	38	1				
	37	0				
	36	0				
	35	2				
	34	3				
Capillary refill time	>4	3				
	2-4	2				
	0-2	0				
Heart rate (bpm)	+60	3	age (years)	Normal range for heart rate		
	+40	2			<1	110-160
	+20	1			1-2	100-150
	normal	0			2-5	95-140
	-20	1			5-12	80-120
	-40	2			>12	60-100
	-60	3				
neuro	A lert=C=1	0				
	V oice =V= 2	1				
	P ain= P = 3	2				
	Unresponsive =I=4	3				
TOTAL PAWS SCORE		0- 24				

## Annexe 2: The Paediatric Early Warning System (PEWS), Tucker et al

Components	0	1	2	3	Score
<b><u>Behavior</u></b>	Playing/ Appropriate	Sleeping	Irritable	Lethargic/ Confused <u>OR</u> Reduced response to pain	
<b><u>Cardiovascular</u></b>	Pink or Capillary refill 1–2 seconds	Pale or Capillary refill 3 seconds	Grey or Capillary refill 4 seconds <u>OR</u> Tachycardia of 20 above normal rate	Grey and mottled or capillary refill 5 seconds or above. <u>OR</u> Tachycardia of 30 above normal rate or bradycardia	
<b><u>Respiratory</u></b>	Within normal parameters, no retractions	> 10 above Normal Parameters, using accessory muscles <u>OR</u> 30+% FiO2 or 3+ liters/min.	<ul style="list-style-type: none"> <li>• &gt; 20 above normal parameters</li> <li>• Retractions. <u>OR</u></li> <li>• 40+% FiO2 or 6+ liters/min.</li> </ul>	<ul style="list-style-type: none"> <li>• 5 below normal Parameters with retractions</li> <li>• Grunting. <u>OR</u></li> <li>• 50% FiO2 or 8+ liters/min.</li> </ul>	
Score 2 extra for ¼ hourly nebulizers or persistent vomiting following surgery					

Adapted with the author's permission from Monaghan, A. (2005). Detecting and managing deterioration in children. *Paediatric Nursing*, 17, 32–35.

With the author's permission from Tucker K.

### Annexe 3: The Bedside paediatric early warning system score (Bedside PEWS), Parshuram et al

		0	1	2	4
<b>Heart rate (bpm)</b>	0 to < 3 months	> 110 and < 150	≥ 150 or ≤ 110	≥ 180 or ≤ 90	≥ 190 or ≤ 80
	3 to < 12 months	> 100 and < 150	≥ 150 or ≤ 100	≥ 170 or ≤ 80	≥ 180 or ≤ 70
	1-4 years	> 90 and < 120	≥ 120 or ≤ 90	≥ 150 or ≤ 70	≥ 170 or ≤ 60
	> 4-12 years	> 70 and < 110	≥ 110 or ≤ 70	≥ 130 or ≤ 60	≥ 150 or ≤ 50
	> 12 years	> 60 and < 100	≥ 100 or ≤ 60	≥ 120 or ≤ 50	≥ 140 or ≤ 40
<b>Systolic blood pressure (mmHg)</b>	0 to < 3 months	> 60 and < 80	≥ 80 or ≤ 60	≥ 100 or ≤ 50	≥ 130 or ≤ 45
	3 to < 12 months	> 80 and < 100	≥ 100 or ≤ 80	≥ 120 or ≤ 70	≥ 150 or ≤ 60
	1 to 4 years	> 90 and < 110	≥ 110 or ≤ 90	≥ 125 or ≤ 75	≥ 160 or ≤ 65
	> 4 to 12 years	> 90 and < 120	≥ 120 or ≤ 90	≥ 140 or ≤ 80	≥ 170 or ≤ 70
	> 80 and < 100	> 100 and < 130	≥ 130 or ≤ 100	≥ 150 or ≤ 85	≥ 190 or ≤ 75
<b>Capillary refill time</b>		< 3 seconds			≥ 3 seconds
<b>Respiratory rate (breaths/minute)</b>	0 to < 3 months	> 29 and < 61	≥ 61 or ≤ 29	≥ 81 or ≤ 19	≥ 91 or ≤ 15
	3 to < 12 months	> 24 or < 51	≥ 51 or ≤ 24	≥ 71 or ≤ 19	≥ 81 or ≤ 15
	1 to 4 years	> 19 or < 41	≥ 41 or ≤ 19	≥ 61 or ≤ 15	≥ 71 or ≤ 12
	> 4 to 12 years	> 19 or < 31	≥ 31 or ≤ 19	≥ 41 or ≤ 14	≥ 51 or ≤ 10
	> 12 years	> 11 or < 17	≥ 17 or ≤ 11	≥ 23 or ≤ 10	≥ 30 or ≤ 9
<b>Respiratory effort</b>		Normal	Mild increase	Moderate increase	Severe increase/any apnoea
<b>Oxygen saturation (%)</b>		> 94	91 to 94	≤ 90	
<b>Oxygen therapy</b>		Room air		Any to < 4 L/minute or < 50%	≥ 4 L/minute or ≥ 50%

## **Annexe 4: Use of Paediatric Early Warning Scores in intermediate care units.**

Use of Paediatric Early Warning Scores in intermediate care units.

Marie Emilie Lampin<sup>a</sup>, MD; Alain Duhamel<sup>b</sup>, PhD; Hélène Béhal<sup>b</sup>, Biostatistician, Morgan Recher<sup>c</sup>, MD; Francis Leclerc<sup>a</sup>, MD, PhD; and Stephane Leteurtre<sup>a</sup>, MD, PhD

**Address correspondence to:** Marie Emilie Lampin, service de réanimation pédiatrique, Centre Hospitalier Universitaire de Lille, F-59000 Lille, France, [[marie-emilie.lampin@chru-lille.fr](mailto:marie-emilie.lampin@chru-lille.fr)], 33 3 20 44 68 84.

**Affiliations:** <sup>a</sup> Univ. Lille, CHU Lille, EA 2694 - Santé Publique : épidémiologie et qualité des soins, Service de réanimation pédiatrique, F-59000 Lille, France

<sup>b</sup> Univ. Lille, CHU Lille, EA 2694 - Santé Publique : épidémiologie et qualité des soins, Service de Biostatistiques, F-59000 Lille, France; <sup>c</sup> CHU Lille, Service de réanimation pédiatrique, F-59000 Lille, France.

Word count :

## Contributorship statement

Dr Lampin coordinated and supervised the data collection, drafted the initial manuscript, reviewed and revised the manuscript.

Prof. Duhamel and Ms Behal designed the data collection instruments, carried out the initial analyses, and critically reviewed the manuscript for important intellectual content.

Dr Recher participated to acquisition of data and critically reviewed the manuscript for important intellectual content.

Prof. Leteurtre and Prof. Leclerc conceptualized and designed the study, supervised the data collection and critically reviewed the manuscript for important intellectual content.

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

## **ABSTRACT**

### **Objective**

Paediatric Early Warning Scores (EWS) were developed to detect deterioration in paediatric wards or emergency departments. The aim of this study was to assess the relationship between three paediatric EWS and clinical deterioration detected by the nurse in paediatric intermediate care units (PImCU).

### **Methods**

This was a prospective, observational, multicenter study at seven French regional hospitals that included all children < 18 years of age. Clinical parameters included in EWS (the Paediatric Advanced Warning Score, the Paediatric Early Warning Score, and the Bedside Paediatric Early Warning System) were prospectively recorded every 8 hours or in case of deterioration. The outcome was a call to physician by the nurse when a clinical deterioration was observed. The cohort was divided into derivation and validation cohorts. An update methodology for repeated measures was used and discrimination was estimated by the area under the ROC curve (AUC).

### **Results**

A total of 2,636 children were included for 14,708 observations to compute a posteriori the EWS. The discrimination of the three EWS for predicting calls to physicians by nurses was good (range: 0.87 to 0.91) for the derivation cohort and moderate (range: 0.71 to 0.76) for the validation cohort. Equations for probability thresholds of calls to physicians, taking into account the time  $t$ , the score at time  $t$ , and the score at admission, are available.

### **Conclusion**

These three EWS developed for children in paediatric wards or emergency departments can be used in PImCU to detect a clinical deterioration and predict the need for medical intervention.

**Keywords:** child, intermediate care units, early warning score, clinical deterioration, performance

## **INTRODUCTION**

Intermediate care units (ImCU) or high dependency care units (HDC) are units between regular wards and intensive care units (ICUs) for patients who require monitoring due to potential organ failure <sup>1</sup>. A review on utilization of ImCU reported diversity in their formats but common denominators are continuous monitoring and respiratory support, without mechanical ventilation and multiple vasoactive medications <sup>2</sup>. Children hospitalized in Paediatric ImCU (PImCU) are at high risk of deterioration. Many paediatric Early Warning Scores (EWS), aimed at the detection of deterioration, have been developed for patients in paediatric wards or emergency departments, using transfer to PICU, request for emergency assistance, or cardiac arrest as outcome variables <sup>3 4</sup>. These physiology-based scoring systems should alert staff to detect deterioration and accelerate access to appropriate intervention. No EWS has been developed for the population of PImCU to predict clinical deterioration. The aim of this study was to assess the relationship between three paediatric EWS (the Paediatric Advanced Warning Score (PAWS) <sup>5</sup>, the Paediatric Early Warning Score (PEWS) <sup>6</sup>, and the Bedside Paediatric Early Warning System (Bedside PEWS) <sup>7</sup> and the deterioration of a child's condition detected by the nurse in PImCU and predict the need for medical intervention.

## **METHODS**

### **Setting**

We performed a prospective observational study in seven PImCU of regional hospitals in northern France, each comprising 4 to 6 beds. Recruitments took place between 7 September 2012 and 7 January 2014.

All patients admitted were assessed for eligibility.

### **Data collection**

Collected data included demographic parameters, medical background, course of the care, primary reason for admission, and primary disease at admission.

Clinical parameters (referred to as “observations” in this paper) were collected by nurses on patient day sheets (standardized for the study in the seven hospitals) for each patient at PImCU admission (=H0), every 8 hours, and at each time the nurse detected deterioration in the child’s condition (collection data form provided in Appendix A, table S1). These parameters included temperature, heart rate, blood pressure, capillary refill time, oxygen saturation, respiratory rate, work of breathing (0 to 3), apnea, oxygen therapy (room air/cannula / mask, flow and FIO<sub>2</sub>), level of consciousness (conscious / voice response / pain response / unconscious). The three EWS, easy to use and routinely feasible at the bedside (PAWS coted 0-21; PEWS 0-9 and Bedside PEWS 0-26; detailed in Appendix A, Table S2, Table S3, Table S4), were computed a posteriori from these clinical parameters; note these were not used to inform the decision-making at the bedside. So, the value of the score was not available to nurses and physicians, and, thus not used to inform them.

### **Outcome**

A call to physician by the nurse (Yes or No) was the outcome variable. It was reported “yes” only when the physician was called because the nurse was worried about the child’s condition.

### **Reproducibility**

One day per week and for one observation per day and per patient (day and time of the collection left to the choice of the teams), the clinical parameters were collected simultaneously by the nurse in charge of the child and by a second observer (the attending physician), to determine reproducibility. The collection time was chosen by the team of the different centers but once chosen, remained the same throughout the study.

### **Sample size**

We estimated the number of calls to physician by nurses in hospitalized patients during their entire stay around 10%, from Tucker et al <sup>6</sup>. Calculation of the sample size was based on the

method described by Hanley JA and McNeil BJ <sup>8</sup>. We assumed an expected Area Under the receiver-operating Curve (AUC) of 0.80. To show that the AUC was strictly better than 0.75, with a one-sided test (type I error= 0.05, power=0.8) required 2,000 patients. We, therefore, sought to recruit 2,500 patients (corresponding to a power greater than 0.85).

### **Statistical analysis**

When one or two of the clinical parameters were missing, the score was calculated by considering the missing values as normal. The scores with more than two missing parameters were not analyzed. The missing data for each score are presented in table S5 (Appendix A). A sensitivity analysis was conducted to determine the impact of default to normal imputation compared to a complete case analysis (these results are presented in table S6, Appendix B). There were no differences between the AUC. The method of default to normal imputation was used up to 2 missing items per score. The interobserver agreement (nurse-physician) was assessed by calculating the intraclass correlation coefficient for each score.

The cohort was divided into derivation (the first 70% included patients) and validation (the remaining 30%) cohorts. In the derivation cohort, a logistic regression model with a random subject effect (General Linear Mixed Model) to account for multiple observations per subject was performed. In this model, the fixed effects were the time  $t$ , the score at time  $t$ , and the score at admission. This model allowed estimation of the predicted probability at each time  $t$ . The discriminant power of the model was estimated by the AUC associated with the predicted probabilities. The odds ratio (OR) of each score adjusted for the score at admission and time was calculated with their 95% confidence interval. For each score, we computed the threshold for the predicted probability, maximizing the specificity for a sensitivity fixed to 90% allowing establishment of a prediction rule for calls to physician by nurses.

In the validation cohort, the AUC were calculated using the coefficients of the fixed effects estimated on the derivation cohort as done by Foulkes et al <sup>9</sup>. We assessed the stability of the

prediction rules by computing the sensitivity and specificity corresponding to the thresholds determined in the derivation cohort. Whilst we choose to exclude a random effect for center from our prediction model, as it is not possible to factor this into a prediction rule for calls to the physician, we provide information on the center differences, and the impact of center variation on the AUC in table S7 and Table S8, Appendix B. The analyses were performed using the observations from admission to 24 hours (as median length of stay), to 36 hours, to 48hours and to 6 days (6 days corresponding to the 95<sup>th</sup> percentiles of the duration of the stay in PImCU).

All analyses were performed using a two-tailed test with an alpha level of 0.05. Statistical analyses were performed using SAS software version 9.4 (SAS Institute, Cary, NC, USA).

Exclusion criteria of the statistical analysis:

Patients who did not have scores at H0 or for whom no score could be computed after H0 were excluded and the observations recorded after 6 days were excluded from analysis.

### **Ethical considerations**

This was an observational study which required no intervention; therefore the institutional review board “Société de Réanimation de Langue Française (SRLF)” waived the need for informed consent. All patients and their parents who were able to provide consent received written and oral information prior to the study and had the option for their data to be excluded from the study. The study and its database were declared to be safe and approved by the French authorities (Commission Nationale de l’Informatique et des Libertés) (DR-2012-594), and by the ethics committee (CE SRLF 12-351).

### **RESULTS**

Of the 2,909 patients assessed for eligibility, 2,636 were included. A flowchart is presented in Figure 1 and patients’ characteristics are listed in Table 1.

Table 1: Baseline characteristics of the paediatric intermediate care unit population

Baseline characteristics	Total,	Included patients	Excluded patients*	p
--------------------------	--------	-------------------	--------------------	---

	n=2,870	n=2,636	n=234	
Sex ratio	1.26	1.26	1.32	ns
Median age months [IQR]	29 [5-103]	28 [5-98]	42 [7-162]	0.01
Median length of stay days [IQR]	1.0 [1.0-3.0]	1.5 [0.8-2.7]	0.2 [0.1-0.3]	< 10 <sup>-4</sup>
Comorbidities, n(%)**	945 (32.9)	876 (33.2)	69 (29.5)	ns
Location prior to admission, n(%)**				0.001
Emergency	2006 (69.9)	1851 (70.2)	155 (66.2)	
Pediatric ward	386 (13.4)	361 (13.7)	25 (10.7)	
Operating room	71 (2.5)	68 (2.6)	3 (1.3)	
PICU	25(0.9)	25 (0.9)	0 (0.0)	
Home	377(13.1)	328 (12.4)	49 (20.9)	
Type of stay, n(%)**				ns
Medical	2714 (94.6)	2490 (94.5)	224 (95.7)	
Surgical	119 (4.1)	113 (4.3)	6 (2.6)	
Primary reason for admission, n(%)**				0.002
Respiratory	1272 (44.3)	1191(45.2)	81(34.6)	
Neurologic	627 (21.8)	580(22.0)	47(20.0)	
Cardiovascular	218 (7.6)	198(7.5)	20(8.5)	
Gastrointestinal	222 (8.2)	202(7.7)	20(8.5)	
Others	529 (18.4)	465 (17.6)	64 (27.3)	
Cause of primary diseases, n(%)**				< 10 <sup>-4</sup>
Infectious	1546 (53.9)	1450 (55.0)	96 (41.0)	
Toxicological	201 (7.0)	176 (6.7)	25 (10.7)	
Congenital	136 (4.7)	127 (4.8)	9 (3.8)	
Trauma	126 (4.4)	114 (4.3)	12 (5.1)	
Oncological	65 (2.3)	42 (1.6)	23 (9.8)	
Others	794 (27.7)	727 (27.6)	67 (28.6)	
Destination n(%)**				
Home	1024 (35.7)	937 (35.5)	87 (37.2)	
Paediatric ward	1586 (55.3)	1487 (56.4)	99 (42.3)	
Others	246 (8.6)	201 (7.6)	45 (19.2)	
Transfer to PICU n(%)**	85 (3.0)	68 (2.6)	17 (7.3)	< 10 <sup>-4</sup>
Death, n	1	1	0	
Median score at admission [IQR]				
PAWS	2 [0-4]	3 [2-4]	2 [1-4]	0.001
PEWS	0 [0-2]	1 [0-3]	0 [0-3]	0.005
Bedside PEWS	3 [0-5]	3 [2-6]	3 [1-5]	0.004

ns: not significant, IQR: interquartile range, PICU: paediatric intensive care unit, PAWS: Paediatric Advanced Warning Score; PEWS: Paediatric Early Warning Score; Bedside PEWS: Bedside Paediatric Early Warning System

\*234 excluded patients: lost records (n=2), at least one score missing at H0 (n=27), at least one score could not be computed after H0 (n=205).

\*\*Numbers and percentages take into account missing data.

The results of the comparison between excluded and included patients (234 versus 2,636 patients) are presented in Table 1. Of the excluded patients, 88% (205/234) were excluded because the three scores could not be calculated at least once after H0, which implies a very short duration of stay (< 8 hours). These patients were older, more frequently at home before admission to the PImCU (20.9%), had a lower incidence of respiratory failure (34.6%) at admission, and a lower rate of infection (41.0%) than the included patients (all p < 0.05).

Reproducibility of the three scores was good: the intraclass correlation coefficients (95% CI) for PAWS, PEWS, and Bedside PEWS were 0.807 (0.802-0.812) (n=372 observations), 0.857 (0.853-0.861) (n=375), and 0.806 (0.801-0.811) (n=367), respectively.

Of 19071 observations collected, 14,708 were included into analysis allowing calculating 12,668 (86.1%) PAWS, 11,756 (79.9%) PEWS, and 12,191 (82.9%) Bedside PEWS complete scores and 1,710 (11.6%) PAWS, 2,916 (19.8%) PEWS, and 2,078 (14.1%) Bedside PEWS additional scores, after considering one or two missing parameters as normal.

A call to physician by nurse occurred in 1,064 (7%) observations. The median scores for calls to physician by nurse were 3 [2-5] for the PAWS, 2 [0-3] for the PEWS, and 5 [2-8] for the Bedside PEWS.

Performances of the three scores for predicting calls to physician by nurse on the derivation cohort are presented in Table 2. These results regarded the complete scores and imputed scores with default to normal up to two missing items: sensitivity analysis results were comparable (Table S6, Appendix B). AUC of the three scores were between 0.87 ( $\leq 24$  hours) and 0.91 ( $\leq 6$  days). Equations of probability thresholds of calls to physician, taking into account the time  $t$ , the score at the  $t$  time, and the score at admission are provided in Table 2.

Table 2: Performances of the PAWS, PEWS, and Bedside PEWS scores on the derivation sample on the first 24 hours, first 36 hours, first 48 hours, and first six days.

Total scores analyzed (complete and imputed with “default to normal” method)					
Scores	AUC (95% CI)	Probability threshold (%) <sup>a</sup>	Specificity (%)	PPV (%)	NPV (%)
$\leq 24$ hours [n=4,693 observations; call to physician = 494 (11%) observations]					
PAWS <sup>b</sup>	0.87 (0.85-0.89)	5.9	49	17	98
PEWS <sup>c</sup>	0.87 (0.86-0.89)	6.0	48	17	98
Bedside PEWS <sup>d</sup>	0.87 (0.85-0.89)	5.2	48	16	98
$\leq 36$ hours [n=5,892 observations; call to physician = 559 (9%) observations]					
PAWS <sup>b</sup>	0.90 (0.88-0.91)	7.3	64	20	98
PEWS <sup>c</sup>	0.89 (0.88-0.91)	6.7	55	17	98
Bedside PEWS <sup>d</sup>	0.89 (0.87-0.91)	6.2	58	17	98
$\leq 48$ hours [n=7,196 observations; call to physician = 629 (9%) observations]					
PAWS <sup>b</sup>	0.91 (0.90-0.92)	7.5	69	22	99

PEWS <sup>c</sup>	0.91 (0.89-0.92)	7.4	71	23	99
Bedside PEWS <sup>d</sup>	0.91(0.89-0.92)	6.7	70	21	99
<hr/>					
$\leq 6$ day [n=10,152 observations; call to physician = 749 (7%) observations]					
PAWS <sup>b</sup>	0.91 (0.90-0.92)	7.1	77	23	99
PEWS <sup>c</sup>	0.91 (0.90-0.92)	7.1	76	23	99
Bedside PEWS <sup>d</sup>	0.91 (0.90-0.92)	6.1	75	21	99

<sup>a</sup>Probability threshold maximizing specificity for sensitivity fixed at 90%: probability computed from the General Linear Mixed Model included the following effects: the score at time  $t$ , the score at admission, and the time  $t$ .

<sup>b</sup>predictive equation for *call to physician*:  $Logit = -2.867 + 0.315*(score\ at\ the\ time\ t) - 0.072*(score\ at\ admission) - 0.016*(time\ t)$ ; Proba of *call to physician* =  $1/(1+\exp(-Logit))$

<sup>c</sup> predictive equation for *call to physician*:  $Logit = -2.630 + 0.448*(score\ at\ the\ time\ t) - 0.101*(score\ at\ admission) - 0.017*(time\ t)$ ; Proba of *call to physician* =  $1/(1+\exp(-Logit))$

<sup>d</sup> predictive equation for *call to physician* :  $Logit = -3.033 + 0.263*(score\ at\ the\ time\ t) - 0.058*(score\ at\ admission) - 0.015*(time\ t)$ ; Proba of *call to physician* =  $1/(1+\exp(-Logit))$

Example:

If PAWS at admission=3, time=48 hours, PAWS (H48)=4

Predictive equation for *call to physician*:  $Logit = -2.867 + 0.315*(4) - 0.072*(3) - 0.016*(48) = -2.591$

Proba of call to physician =  $1/(1+e^{-Logit}) = 7.0\%$

This Probability is less than 7.5% (probability threshold for PAWS from admission to 48h): so nurse must not call physician.

*Abbreviations*: AUC: Area Under the receiver-operating Curve; CI: confidence interval; PAWS: Paediatric Advanced Warning Score; PEWS: Paediatric Early Warning Score; Bedside PEWS: Bedside Paediatric Early Warning System; PPV: positive predictive value; NPV: negative predictive value.

Specificities were 48% and 49% ( $\leq 24$  hours) and increased to 75 and 76% ( $\leq 6$  days) (Table 2). Negative predictive values were excellent (98 and 99%) for all scores and at all time points (Table 2).

Each one-point increase in the PAWS, PEWS, and Bedside PEWS significantly increased the risk of a physician being called (OR [95% CI]: 1.37 [1.30-1.44], 1.57 [1.47-1.67], and 1.30 [1.26-1.35], respectively).

The prediction rule built from the derivation cohort (if the predicted probability was greater than the threshold) was applied to the validation cohort (Table 3).

Table 3: Performances of the PAWS, PEWS, and Bedside PEWS scores on the validation sample on the first 24 hours, first 36 hours, first 48 hours, and first six days.

	AUC (95% CI)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
<hr/>					
$\leq 24$ hours [n=2,028 observations; call to physician = 219 (11%) observations]					
PAWS	0.74 (0.70-0.78)	76	50	15	95
PEWS	0.74 (0.70-0.78)	77	48	15	95
Bedside PEWS	0.76 (0.72-0.80)	81	46	15	95
<hr/>					
$\leq 36$ hours [n=2,550 observations; physician call = 245 (10%) observations]					
PAWS	0.74 (0.70-0.77)	72	60	16	95

PEWS	0.75 (0.71-0.78)	78	53	15	96
Bedside PEWS	0.75 (0.72-0.79)	79	53	15	96
<hr/>					
$\leq 48$ hours [n=3,156 observations; call to physician = 267 (8%) observations]					
PAWS	0.73 (0.70-0.77)	64	71	17	96
PEWS	0.75 (0.72-0.79)	73	63	15	96
Bedside PEWS	0.75 (0.72-0.79)	71	65	15	96
<hr/>					
$\leq 6$ day [n=4,556 observations; call to physician = 315 (7%) observations]					
PAWS	0.71 (0.68-0.74)	55	76	14	96
PEWS	0.75 (0.72-0.78)	58	75	15	96
Bedside PEWS	0.75 (0.72-0.78)	66	69	13	96

Performances of scores were calculated from a predictive equation derived on the development sample; sensitivity and specificity were calculated using thresholds measured on the development sample.

*Abbreviations:* AUC: Area Under the receiver-operating Curve; CI: confidence interval; PAWS: Paediatric Advanced Warning Score; PEWS: Paediatric Early Warning Score; Bedside PEWS: Bedside Paediatric Early Warning System, AUC: area under the curve, PPV: positive predictive value, NPV: negative predictive value.

AUC of the three scores were between 0.71 (PAWS  $\leq 6$  days) and 0.76 (Bedside PEWS  $\leq 24$  hours). Sensitivities of the three scores were maximal at H24 (76 to 81%) and decreased until day-6 (55% to 66%) while the prevalence of calls to physician also decreased from 11 to 7%. Negative predictive values were excellent (95 and 96%) for all scores and at all-time points.

## DISCUSSION

This study is the first to assess the use of paediatric EWS in PImCU in detecting deterioration of children. The discriminative ability estimated by the AUCs of the three scores for predicting calls to physician were good to excellent (range: 0.87 to 0.91) on the derivation cohort and were moderate (range: 0.71 to 0.76) on the validation cohort. Equations of probabilities thresholds of calls to physician, taking into account the score at time  $t$ , the score at admission, and time  $t$ , are available. The three EWSs had good interobserver reproducibilities.

PImCU are recent units in which children at high risk of deterioration are hospitalized<sup>1</sup>. Four systematic reviews of paediatric EWS have been published<sup>3,4,10,11</sup>. To develop or validate paediatric EWS, a gold standard that establishes clinical deterioration is necessary, but there

is no consensus about the gold standard for this event<sup>3</sup>. There were multiple outcome measures: death<sup>12</sup>, cardiac arrest or code blue<sup>13-15</sup>, unplanned transfer to PICU or requirement for PICU<sup>6,7,15-21</sup>, a call for urgent medical assistance or rapid response system (RRS) activation<sup>22-25</sup>, and length of hospital stay<sup>22</sup>. However, the aim of these EWS was to identify deterioration before respiratory or cardiac arrest or transfer to a PICU. Moreover, most of these outcomes are rare events, and this affects the methodology used for the validation of EWS. In our study, the outcome variable used was call to physician by nurses in case of deterioration. This outcome is not ideal as it is influenced by numerous factors: experience of the nurse, knowledge of the patient, relationship between the nurse and the physician, ease of calling and workload. Call to physician by nurses may be a confounding factor because the collect of clinical parameter at the time of “deterioration” depends on subjective clinical judgment by nurse. It was not verified if call to physician by the nurse was justified. Bonafide et al, in qualitative study to identify mechanisms beyond the statistical ability of use EWS by nurses and physician suggested that combination of EWS and clinical judgment could be a better system for detecting deterioration<sup>26</sup>. Jensen et al, in a study interested in factors that may compromise the use of EWS in clinical practice reported also the lack of clinical judgment in EWS<sup>27</sup>. There is no gold standard to detect clinical deterioration and other outcomes used such as death, cardiac arrest or code blue, unplanned transfer to PICU or requirement for PICU are too rare events to be used in a prospective methodology including all consecutive patients. Many EWS were retrospectively developed and their validity was evaluated using a case/control type of methodology with a number of patients of less than 120<sup>14-16,19,20,28,29</sup>. Our study was large prospective and multicentric, with clinical parameters of the scores (computed a posteriori) recorded for each patient at admission, every 8 hours, and in case of deterioration and outcome (call to physician by the nurse) evaluated for all collected scores. So there is a dependent relationship between regular

observations evaluated by the nurse either every 8 hours either or at any time of the stay. In current practice, a score to detect deterioration that is repeated at regular intervals is more relevant than a single score. We have used an update methodology for repeated measures. However, a score threshold could not be identified in this mixed model due to repeated measures in the same patient. The thresholds that have been proposed for EWS are varied and can not compared each other<sup>4</sup>. Using the outcome *cardiac arrest*, *RRS activation*, or *PICU admission*, in case /control studies, thresholds proposed for the Bedside PEWS were  $\geq 7$  and  $\geq 8$ <sup>7,14,16,17,28</sup>. For different modified PEWS, the proposed thresholds were between  $\geq 2$  and  $\geq 5$ <sup>6,13,14,18,19,29</sup> and for PAWS  $\geq 3$ <sup>5</sup>. Two prospective studies used *PICU admission* as the outcome: the threshold was  $\geq 1$  for PEWS<sup>30</sup> and  $\geq 3$  for the PAWS<sup>31</sup>. Sieger et al. analyzed the validity of different EWS's in 17,943 children and observed that the optimal threshold to calculate sensitivity and specificity for PICU admission was low (threshold at 1), except for the Bedside PEWS, which had a threshold of 3: AUCs were 0.77 for the PAWS, 0.79 for the PEWS and 0.82 for the Bedside PEWS<sup>32</sup>. In our study AUCs of the three scores were good to excellent (0.87-0.91) for the derivation cohorts and moderate (0.71-0.76) for the validation cohorts. On validation cohort, sensitivity decreased (from 76-81% to 55-66%) and specificity increased over time for the three scores (from 46-50% to 69-76). Mulherin et al, indicated that sensitivity and specificity subgroup variation were not a bias but is clinically relevant information to be identified and reported<sup>33</sup>. These authors suggested replacing the term "spectral bias" by "spectrum effect" that reflected the heterogeneity in the test performance when applied to different subgroups, like in our study, with different subgroups depending on length of PImCU stay. These three scores seem better detect deterioration at the beginning of the stay and their performances decreased over time; this may be related to the stabilization of patients, thus requiring fewer calls to physician throughout the stay.

The reproducibility of the three scores was good (coefficient range: 0.81 to 0.86). Few studies have analyzed the interobserver reproducibility of these EWS: Chaiyakulsil et al<sup>31</sup> reported a good inter-rater reliability (kappa = 0.75) for the PAWS and Gold et al<sup>30</sup> reported an excellent inter-rater reliability (intraclass coefficient = 0.91) for the PEWS.

None of the three scores appeared to be better than the others for detecting deterioration. The PAWS may be preferred for quick initial assessment in emergency care or in the ward because blood pressure measurement is not necessary. The Bedside PEWS has the advantage of considering oxygen saturation and oxygen supply, contrary to the PAWS, which only considers oxygen saturation, but it does not include any neurological assessment. The PEWS, is divided into three categories - respiratory, circulatory, and neurological, corresponding to the different types of organ failure- and, thus, seems easier to use.

Sambecq et al, in a cross-sectional survey, revealed that the use of 45 different EWS scores in Dutch hospitals, can lead to a false sense of security and recommended to establish a national working group to coordinate implementation of a EWS usable for both general and university hospitals<sup>34</sup>. A recent review emphasized that despite widespread use, the evidence base for EWS remains limited because there is no consensus of the most effective EWS and it exists a lack of robust, valid and clinically meaningful outcomes<sup>35</sup>. In the EPOCH Randomized Clinical Trial, including 144 539 patients, the effect of the BedsidePEWS intervention was assessed on all-cause hospital mortality and late admission to PICU, cardiac arrest and ICU resource use. Despite this large number of patients, to compensate the low number of events/outcome, responsible for the lack of robustness on the validity of the EWS, the BedsidePEWS intervention did not significantly decrease all-cause mortality, compared with usual care<sup>36</sup>. Some studies have proposed other approaches to test pediatrics EWS: Jensen et al, in a multicenter, randomized controlled trial, comparing two different EWS models to predict deterioration requiring transfer to a higher level of care<sup>37</sup>. In this study too,

despite a large number of enrolled patients (n=16213), 22 unplanned transfers to a higher level of care were identified. No significant difference between the two scores (Bedside PEWS and CDR PEWS) was identified but CDR PEWS seemed more acceptable to staff<sup>37</sup>. Thomas-Jones et al, proposed a prospective, mixed-methods, before and after study based approach to improvement. This study is still ongoing<sup>38</sup>. We propose another approach: the use of a prediction rule taking into account the EWS on admission and at the time t when the deterioration occurs, rather than a score threshold or an escalation algorithm indicating the care-team action<sup>36</sup>.

### Limitations

First, there are missing data. The score was calculated considering the missing values (1 or 2) as normal. How are managed missing data is not often explained. For the Bedside PEWS, Parshuram et al, took the most recent recorded data when the corresponding data was missing<sup>17</sup>. Some authors attributed normal values whatever the number of missing parameters<sup>5,14,15</sup>. Others used a multiple imputation model<sup>31,32</sup>. In our study, multiple imputation was not used because this would have added repetition in a model in which there were already repeated measurements. Second, although this study was multicentric, it included seven PImCU of regional hospitals from the same region, which may represent a recruitment bias. Patients between centers could be different. However, the potential center effect could not be added as random effect in the model because it would not allow to obtain a rule of predictive decision of call to physician by the nurse. Furthermore, when we adjusted to the center, the discriminant power evaluated by the AUC was similar. Moreover, the patients included in our study were admitted to PImCU not attached in PICU and these patients may be different (maybe less severe) from the patients hospitalized in PImCU attached to PICU. As reported by Plate et al, there is a great diversity of these ImCU<sup>2</sup>. A focus on High Dependency Care for Children in UK recommends to separate three care level for activity of these HDC (level

1: enhanced care unit, level 2: critical care unit, level 3: intensive care unit). Our population would correspond to level 1 or 2<sup>39</sup>. This could explain our transfer rate to PICU (3%), barely higher than that reported by Sieger et al in an emergency department (2%)<sup>32</sup>. Third, because we used a methodology for repeated measures, we do not provide an absolute threshold value but equations of probability thresholds for calls to a physician, taking into account the score at time  $t$ , the score at admission and time  $t$  are available.

## CONCLUSIONS

This prospective multicenter observational study indicates that three EWS (PAWS, PEWS, and Bed PEWS) initially developed for children admitted to paediatric wards or presenting to the emergency departments can be used in PImCU to predict the need for medical intervention. Further studies are needed in different contexts and different countries to validate the usefulness EWS with prediction rule for call to physician in case of deterioration.

**Acknowledgments** The authors would like to thank nurses from seven regional hospitals and Isabelle Grit for her logistic help.

**Competing Interests** None declared

**Funding Source:** This study was supported by a grant from the French Ministry of Health (PHRC 2011). The funding agency did not have a role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**What is known about this topic:** No EWS have been tested in PImCU while these children are at high risk of clinical deterioration.

**What This Study Adds:** It is the first study to focus on EWS used in PImCU and proposes call to physician by nurse as outcome for deterioration in patients.

## Legends:

Figure 1: Flowchart of the inclusions

\*The distribution of patient inclusions in the seven centers were 369 for center 1, 333 for center 2, 477 for center 3, 246 for center 4, 272 for center 5, 465 for center 6, and 474 for center 7.

## REFERENCES

1. Jaimovich DG, American Academy of Pediatrics Committee on Hospital Care and Section on Critical Care. Admission and discharge guidelines for the pediatric patient requiring intermediate care. *Pediatrics* 2004;113:1430-3.
2. Plate JDJ, Leenen, LPH., Houwert, M., et al. Utilisation of Intermediate Care Units: A Systematic Review. *Crit Care Res Pract* 2017;doi:10.1155/2017/8038460
3. Lambert V, Matthews A, MacDonell R, et al. Paediatric early warning systems for detecting and responding to clinical deterioration in children: a systematic review. *BMJ Open*. 2017;7:e014497.
4. Chapman SM, Wray J, Oulton K, et al. Systematic review of paediatric track and trigger systems for hospitalised children. *Resuscitation* 2016;109:87-109.
5. Egdell P, Finlay L, Pedley DK. The PAWS score: validation of an early warning scoring system for the initial assessment of children in the emergency department. *Emerg Med J* 2008;25:745-9.
6. Tucker KM, Brewer TL, Baker RB, et al. Prospective evaluation of a pediatric inpatient early warning scoring system. *J Spec Pediatr Nurs* 2009;14:79-85.
7. Parshuram CS, Hutchison J, Middaugh K. Development and initial validation of the Bedside Paediatric Early Warning System score. *Crit Care Lond Engl* 2009;13:R135.
8. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29-36.  
doi:10.1148/radiology.143.1.7063747
9. Foulkes AS, Azzoni L, Li X, et al. Prediction based classification for longitudinal biomarkers. *Ann Appl Stat* 2010;4:1476-97.
10. Chapman SM, Grocott MPW, Franck LS. Systematic review of paediatric alert criteria for identifying hospitalised children at risk of critical deterioration. *Intensive Care Med* 2010;36:600-11.
11. Trubey R, Huang C, Lugg-Widger FV, et al. Validity and effectiveness of paediatric early warning systems and track and trigger tools for identifying and reducing clinical deterioration in hospitalised children: a systematic review. *BMJ Open* 2019;9:e022105.
12. Olson D, Davis NL, Milazi R, et al. Development of a severity of illness scoring system (inpatient triage, assessment and treatment) for resource-constrained hospitals in developing countries. *Trop Med Int Health* 2013;18:871-8.
13. Akre M, Finkelstein M, Erickson M, et al. Sensitivity of the pediatric early warning score to identify patient deterioration. *Pediatrics* 2010;125:e763-769.
14. Robson M-AJ, Cooper CL, Medicus LA, et al. Comparison of three acute care pediatric early warning scoring tools. *J Pediatr Nurs* 2013;28:e33-41.
15. Duncan H, Hutchison J, Parshuram CS. The Pediatric Early Warning System score: a severity of illness score to predict urgent medical need in hospitalized children. *J Crit Care* 2006;21:271-8.

16. Parshuram CS, Duncan HP, Joffe AR, et al. Multicentre validation of the bedside paediatric early warning system score: a severity of illness score to detect evolving critical illness in hospitalised children. *Crit Care Lond Engl* 2011;15:R184.
17. Fuijkschot J, Vernhout B, Lemson J, et al. Validation of a Paediatric Early Warning Score: first results and implications of usage. *Eur J Pediatr* 2015;174:15-21.
18. Mandell IM, Bynum F, Marshall L, et al. Pediatric Early Warning Score and unplanned readmission to the pediatric intensive care unit. *J Crit Care* 2015;30:1090-5.
19. Agulnik A, Forbes PW, Stenquist N, et al. Validation of a Pediatric Early Warning Score in Hospitalized Pediatric Oncology and Hematopoietic Stem Cell Transplant Patients. *Pediatr Crit Care Med* 2016;17:e146-53.
20. Skaletzky SM, Raszynski A, Totapally BR. Validation of a modified pediatric early warning system score: a retrospective case-control study. *Clin Pediatr (Phila)* 2012;51:431-5.
21. Tume L. The deterioration of children in ward areas in a specialist children's hospital. *Nurs Crit Care* 2007;12:12-9.
22. Bell D, Mac A, Ochoa Y, et al. The Texas Children's Hospital Pediatric Advanced Warning Score as a predictor of clinical deterioration in hospitalized infants and children: a modification of the PEWS tool. *J Pediatr Nurs* 2013;28:e2-9.
23. Parshuram CS, Bayliss A, Reimer J, et al. Implementing the Bedside Paediatric Early Warning System in a community hospital: A prospective observational study. *Paediatr Child Health* 2011;16:e18-22.
24. McKay H, Mitchell IA, Sinn K, et al. Effect of a multifaceted intervention on documentation of vital signs and staff communication regarding deteriorating paediatric patients. *J Paediatr Child Health* 2013;49:48-56.
25. Randhawa S, Roberts-Turner R, et al. Implementing and sustaining evidence-based nursing practice to reduce pediatric cardiopulmonary arrest. *West J Nurs Res* 2011;33:443-56.
26. Bonafide CP, Roberts KE, Weirich CM, et al. Beyond statistical prediction: qualitative evaluation of the mechanisms by which pediatric early warning scores impact patient safety. *J Hosp Med* 2013;8:248-53.
27. Jensen CS, Nielsen PB, Olesen HV, et al. Pediatric Early Warning Score Systems, Nurses Perspective - A Focus Group Study. *J Pediatr Nurs* 2018;41:e16-22.
28. Gawronski O, Ciofi Degli Atti ML, Di Ciommo V, et al. Accuracy of Bedside Paediatric Early Warning System (BedsidePEWS) in a Pediatric Stem Cell Transplant Unit. *J Pediatr Oncol Nurs* 2016;33:249-56.
29. McLellan MC, Gauvreau K, Connor JA. Validation of the Cardiac Children's Hospital Early Warning Score: an early warning scoring tool to prevent cardiopulmonary arrests in children with heart disease. *Congenit Heart Dis* 2014;9:194-202.

30. Gold DL, Mihalov LK, Cohen DM. Evaluating the Pediatric Early Warning Score (PEWS) system for admitted patients in the pediatric emergency department. *Acad Emerg Med* 2014;21:1249-56.
31. Chaiyakulsil C, Pandee U. Validation of pediatric early warning score in pediatric emergency department. *Pediatr Int* 2015;57:694-8.
32. Seiger N, Maconochie I, Oostenbrink R, et al. Validity of different pediatric early warning scores in the emergency department. *Pediatrics* 2013;132:e841-50.
33. Mulherin SA, Miller WC. Spectrum bias or spectrum effect? Subgroup variation in diagnostic test evaluation. *Ann Intern Med* 2002;137:598-602.
34. Sambeek SJ van, Fuijkschot J, Kramer BW, et al. Pediatric Early Warning System Scores: Lessons to be Learned. *J Pediatr Intensive Care* 2018;7:27-32.
35. Chapman SM, Maconochie IK. Early warning scores in paediatrics: an overview. *Arch Dis Child*. 2019;104:395-9.
36. Parshuram CS, Dryden-Palmer K, Farrell C, et al. Effect of a Pediatric Early Warning System on All-Cause Mortality in Hospitalized Pediatric Patients: The EPOCH Randomized Clinical Trial. *JAMA* 2018;319:1002-12.
37. Jensen CS, Olesen HV, Aagaard H, et al. Comparison of Two Pediatric Early Warning Systems: A Randomized Trial. *J Pediatr Nurs* 2019;44:e58-65.
38. Thomas-Jones E, Lloyd A, Roland D, et al. A prospective, mixed-methods, before and after study to identify the evidence base for the core components of an effective Paediatric Early Warning System and the development of an implementation package containing those core recommendations for use in the UK: Paediatric early warning system - utilisation and mortality avoidance- the PUMA study protocol. *BMC Pediatr*. 2018;18:244.
39. High dependency care for children - time to move on (2014). RCPCH. <https://www.rcpch.ac.uk/resources/high-dependency-care-children-time-move> (accessed July 11, 2019).

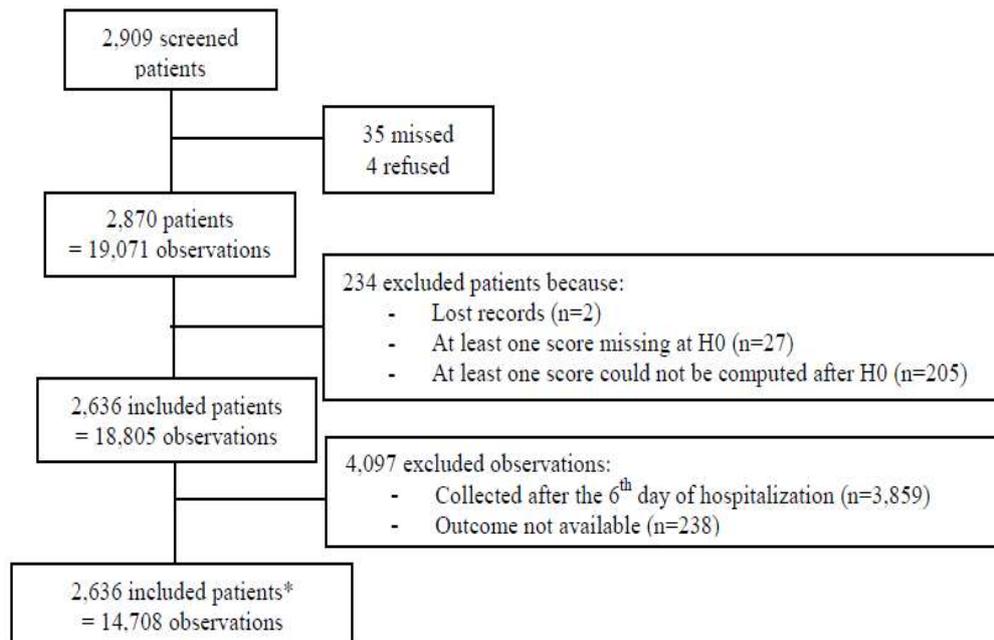


Figure 1: Flowchart of the inclusions

\*The distribution of patient inclusions in the seven centers were 369 for center 1, 333 for center 2, 477 for center 3, 246 for center 4, 272 for center 5, 465 for center 6, and 474 for center 7.



Table S2: PAWS: The Paediatric advanced warning score, Egdell et al <sup>1</sup>

<b>Respiratory rate</b>	+30	3	age (years)	Normal range for respiratory rate		
	+20	2			<1	30-40
	+10	1			1-2	25-35
	normale	0			2-5	25-30
	-10	2			5-12	20-25
					>12	15-20
<b>Work of breathing</b>	Severe increase	3				
	Moderate increase	2				
	normal	0				
<b>Oxygen saturation (%)</b>	>ou= 93%	0				
	90-92	1				
	85-89	2				
	<85	3				
<b>Temprature (°C)</b>	39	2				
	38	1				
	37	0				
	36	0				
	35	2				
	34	3				
<b>Capillary refill time</b>	>4	3				
	2-4	2				
	0-2	0				
<b>Heart rate (bpm)</b>	+60	3	age (years)	Normal range for heart rate		
	+40	2			<1	110-160
	+20	1			1-2	100-150
	normal	0			2-5	95-140
	-20	1			5-12	80-120
	-40	2			>12	60-100
	-60	3				
<b>neuro</b>	A lert=C=1	0				
	V oice =V= 2	1				
	P ain= P = 3	2				
	Unresponsive =I=4	3				
<b>TOTAL PAWS SCORE</b>		<b>0-24</b>				

Table S3: PEWS: The Paediatric Early Warning System, Tucker et al <sup>2</sup>

Components	0	1	2	3	Score
<b><u>Behavior</u></b>	Playing/ Appropriate	Sleeping	Irritable	Lethargic/ Confused OR Reduced response to pain	
<b><u>Cardiovascular</u></b>	Pink or Capillary refill 1–2 seconds	Pale or Capillary refill 3 seconds	Grey or Capillary refill 4 seconds OR Tachycardia of 20 above normal rate	Grey and mottled or capillary refill 5 seconds or above. OR Tachycardia of 30 above normal rate or bradycardia	
<b><u>Respiratory</u></b>	Within normal parameters, no retractions	> 10 above Normal Parameters, using accessory muscles OR 30+% FiO2 or 3+ liters/min.	<ul style="list-style-type: none"> <li>• &gt; 20 above normal parameters</li> <li>• Retractions.</li> <li>OR</li> <li>• 40+% FiO2 or 6+ liters/min.</li> </ul>	<ul style="list-style-type: none"> <li>• 5 below normal Parameters with retractions</li> <li>• Grunting. OR</li> <li>• 50% FiO2 or 8+ liters/min.</li> </ul>	
Score 2 extra for ¼ hourly nebulizers or persistent vomiting following surgery					

Adapted with the author's permission from Monaghan, A. (2005). Detecting and managing deterioration in children. *Paediatric Nursing*, 17, 32–35.

With the author's permission from Tucker K.

Table S4: Bedside PEWS: The Bedside paediatric early warning system score, Parshuram et al<sup>3</sup>

		0	1	2	4
<b>Heart rate (bpm)</b>	0 to < 3 months	> 110 and < 150	≥ 150 or ≤ 110	≥ 180 or ≤ 90	≥ 190 or ≤ 80
	3 to < 12 months	> 100 and < 150	≥ 150 or ≤ 100	≥ 170 or ≤ 80	≥ 180 or ≤ 70
	1-4 years	> 90 and < 120	≥ 120 or ≤ 90	≥ 150 or ≤ 70	≥ 170 or ≤ 60
	> 4-12 years	> 70 and < 110	≥ 110 or ≤ 70	≥ 130 or ≤ 60	≥ 150 or ≤ 50
	> 12 years	> 60 and < 100	≥ 100 or ≤ 60	≥ 120 or ≤ 50	≥ 140 or ≤ 40
<b>Systolic blood pressure (mmHg)</b>	0 to < 3 months	> 60 and < 80	≥ 80 or ≤ 60	≥ 100 or ≤ 50	≥ 130 or ≤ 45
	3 to < 12 months	> 80 and < 100	≥ 100 or ≤ 80	≥ 120 or ≤ 70	≥ 150 or ≤ 60
	1 to 4 years	> 90 and < 110	≥ 110 or ≤ 90	≥ 125 or ≤ 75	≥ 160 or ≤ 65
	> 4 to 12 years	> 90 and < 120	≥ 120 or ≤ 90	≥ 140 or ≤ 80	≥ 170 or ≤ 70
	> 80 and < 100	> 100 and < 130	≥ 130 or ≤ 100	≥ 150 or ≤ 85	≥ 190 or ≤ 75
<b>Capillary refill time</b>		< 3 seconds			≥ 3 seconds
<b>Respiratory rate (breaths/minute)</b>	0 to < 3 months	> 29 and < 61	≥ 61 or ≤ 29	≥ 81 or ≤ 19	≥ 91 or ≤ 15
	3 to < 12 months	> 24 or < 51	≥ 51 or ≤ 24	≥ 71 or ≤ 19	≥ 81 or ≤ 15
	1 to 4 years	> 19 or < 41	≥ 41 or ≤ 19	≥ 61 or ≤ 15	≥ 71 or ≤ 12
	> 4 to 12 years	> 19 or < 31	≥ 31 or ≤ 19	≥ 41 or ≤ 14	≥ 51 or ≤ 10
	> 12 years	> 11 or < 17	≥ 17 or ≤ 11	≥ 23 or ≤ 10	≥ 30 or ≤ 9
<b>Respiratory effort</b>		Normal	Mild increase	Moderate increase	Severe increase/any apnoea
<b>Oxygen saturation (%)</b>		> 94	91 to 94	≤ 90	
<b>Oxygen therapy</b>		Room air		Any to < 4 L/minute or < 50%	≥ 4 L/minute or ≥ 50%

Table S5: Characteristics and missing data of the three scores

Characteristics of the three scores			
	PAWS	PEWS	Bedside PEWS
Item number	7	6	7
Minimum-Maximum	0-21	0-9	0-26
Missing item of the three scores			
	N1/N2	N1/N2	N1/N2
Cardiovascular <sup>1</sup>	NA	888/924 (96.1%)	
Heart rate	27/177 (15.2%)	NA	25/177(14.1%)
Systolic blood pressure	NA	NA	1001/1399 (71.5%)
Capillary refill time	938/1263 (74.3%)	NA	833/1263 (65.9%)
Respiratory <sup>2</sup>	NA	2119/2155(98.3%)	NA
Respiratory rate	275/463 (59.4%)	NA	260/463 (56.1%)
Respiratory effort	192/496 (38.7%)	NA	115/496 (23.2%)
Oxygen saturation	65/243 (26.7%)	NA	62/243 (25.5%)
Oxygen therapy <sup>3</sup>	NA	NA	282/673 (41.9%)
Behavior	64/322 (19.9%)	286/322 (88.8%)	NA
Température	422/652 (64.7%)	NA	NA
Total N1/N2 of each scores	1983/3616 (53.6%)	3293/3401 (96.8%)	2578/4738 (54.4%)
Total observations for the calculation of the three scores, n=14708			
Total imputed scores	1710 (11.6%)	2916 (19.8%)	2078 (14.1%)
Total complete scores	12668 (86.1%)	11756 (79.9%)	12191 (82.9%)
Total analyzed scores	14378 (97.8%)	14672 (99.8%)	14269 (97%)
Total not analyzed scores	330 (2.2%)	36 (0.2%)	439 (3%)

PAWS: Paediatric Advanced Warning Score; PEWS: Paediatric Early Warning Score; Bedside PEWS: Bedside Paediatric Early Warning System  
 N1/N2: number of imputed items with “default to normal” method/ total number of missing items, NA: items not included in the score

<sup>1</sup> Cardiovascular= Capillary refill time OR heart rate for PEWS; <sup>2</sup> Respiratory: Respiratory rate AND Respiratory effort OR Oxygen Flow/FiO2 for PEWS; <sup>3</sup> Oxygen therapy: room air / oxygen flow OR FiO2 in Bedside PEWS.

## Appendix B: Additional statistical analyses

Table S6: Sensitivity analyses to determine the impact of default to normal imputation compared to a complete case analyses

Scores complete and imputed with “default to normal” method		Complete scores
AUC (95% CI)		AUC (95% CI)
Scores		
≤ 24 hours	n=4,693 observations; Call to physician = 494 (11%)	n=4,127 observations Call to physician = 356 (9%)
PAWS <sup>b</sup>	0.87 (0.85-0.89)	0.88 (0.86-0.90)
PEWS <sup>c</sup>	0.87 (0.86-0.89)	0.87 (0.85-0.90)
Bedside PEWS <sup>d</sup>	0.87 (0.85-0.89)	0.85 (0.83-0.88)
≤ 36 hours	n=5,892 observations; Call to physician =559 (9%)	n=5,139 observations Call to physician = 400 (8%)
PAWS <sup>b</sup>	0.90 (0.88-.091)	0.89 (0.87-0.91)
PEWS <sup>c</sup>	0.89 (0.88-0.91)	0.90 (0.88-0.91)
Bedside PEWS <sup>d</sup>	0.89 (0.87-0.91)	0.88 (0.85-0.90)
≤ 48 hours	n=7,196 observations; Call to physician = 629 (9%)	n= 6,284 observations Call to physician =457 (7%)
PAWS <sup>b</sup>	0.91 (0.90-0.92)	0.90 (0.88-0.92)
PEWS <sup>c</sup>	0.91 (0.89-0.92)	0.91 (0.89-0.82)
Bedside PEWS <sup>d</sup>	0.91(0.89-0.92)	0.89 (0.87-0.91)
≤ 6 day	n=10,152 observations; Call to physician = 749 (7%)	n= 8,848 observations Call to physician =562 (6%)
PAWS <sup>b</sup>	0.91 (0.90-0.92)	0.91 (0.90-0.92)
PEWS <sup>c</sup>	0.91 (0.90-0.92)	0.91 (0.89-0.92)
Bedside PEWS <sup>d</sup>	0.91 (0.90-0.92)	0.91 (0.89-0.92)

AUC: Area Under the *receiver-operating* Curve; CI: confidence interval; PAWS: **Paediatric Advanced Warning Score**; PEWS: Paediatric Early Warning Score; Bedside PEWS: **Bedside Paediatric Early Warning System**.

Table S7: Comparison of patient characteristics between the seven centers

	Center 1	Centre 2	Center 3	Center 4	Centre 5	Center 6	Centre 7	p
Median age months [IQR]	14 [4-56]	17[4-65]	53[18-159]	20[5-62]	92[19-181]	24[3-65]	24[3-90]	< 10 <sup>-4</sup>
Median length of stay days [IQR]	1.8[1.0-3.2]	1.4[0.8-2.0]	1.2[0.8-2.0]	1.5[0.7-2.5]	1.6[0.8-2.9]	2.4[1.0-3.8]	0.9[0.7-1.9]	< 10 <sup>-4</sup>
Comorbidities, n(%)	98 (27)	74(22)	180 (38)	69 (28)	127 (47)	158 (34)	170 (36)	< 10 <sup>-4</sup>
Call to physician by nurse	91 (4.0)	55 (3.5)	264 (13.3)	53 (4.3)	254 (14.7)	118 (3.2)	229 (10.1)	< 10 <sup>-4</sup>
Transfer to PICU n(%)	14 (3.8)	3 (0.9)	6 (1.3)	3 (1.2)	9 (3.3)	24 (5.2)	9 (1.9)	4.10 <sup>-4</sup>
Median score at admission [IQR]								
PAWS	4 [2-5]	3[2-5]	3[2-4]	2[1-4]	2[0-4]	3[2-5]	2[1-4]	< 10 <sup>-4</sup>
PEWS	2[0-3]	1[0-3]	1[0-3]	1[0-3]	0[0-2]	2[0-3]	2[1-4]	< 10 <sup>-4</sup>
Bedside PEWS	4[2-7]	3[2-6]	3[1-6]	3[1-6]	2[1-4]	4[2-8]	3[2-6]	< 10 <sup>-4</sup>

, IQR: interquartile range, PICU: paediatric intensive care unit, PAWS: **Paediatric Advanced Warning Score**; PEWS: Paediatric Early Warning Score; Bedside PEWS: **Bedside Paediatric Early Warning System**

Table S8: Effect of center adjustment on AUC

	with adjustment on center analyses AUC (95%CI)	without adjustment on center analyses AUC (95%CI)
PAWS	0.899 (0.888-0.910)	0.914 (0.903-0.925)
PEWS	0.899 (0.887-0.910)	0.910 (0.899-0.921)
Bedside PEWS	0.897 (0.884-0.911)	0.910 (0.899-0.921)

AUC: Area Under the *receiver-operating* Curve; CI: confidence interval; PAWS: **Paediatric Advanced Warning Score**; PEWS: Paediatric Early Warning Score; Bedside PEWS: **Bedside Paediatric Early Warning System**

## References

1. Egdell P, Finlay L, Pedley DK. The PAWS score: validation of an early warning scoring system for the initial assessment of children in the emergency department. *Emerg Med J* 2008;25:745-9.
2. Tucker KM, Brewer TL, Baker RB, et al. Prospective evaluation of a pediatric inpatient early warning scoring system. *J Spec Pediatr Nurs* 2009;14:79-85.
3. Parshuram CS, Duncan HP, Joffe AR, et al. Multicentre validation of the bedside paediatric early warning system score: a severity of illness score to detect evolving critical illness in hospitalised children. *Crit Care Lond Engl* 2011;15:R184.

## Annexe 5: Fiche de recueil des caractéristiques générales

<b>Identité</b> <i>(coller une étiquette nominative avant anonymisation)</i>	Nom de code en 4 signes	__ __ __ __
	Date de naissance : JJ/MM/AA	__ / __ / __
	Date et heure d'admission : JJ/MM/AA (si non renseignés sur l'étiquette)	__ / __ / __ , __ h __
	Sexe	1 Masculin : 2 Féminin I __ I

### Mode d'entrée: choix simple

- |   |  |        |
|---|--|--------|
| 1 <input type="checkbox"/> transfert du service des urgences du même hôpital    | 4 <input type="checkbox"/> admission directe SANS passage par les urgences ou autres services    | I __ I |
| 2 <input type="checkbox"/> transfert post opératoire du même hôpital            | 5 <input type="checkbox"/> transfert de la réanimation pédiatrique ou néonatale du CHRU de Lille |        |
| 3 <input type="checkbox"/> transfert d'un service du même hôpital (sauf 1 et 2) | 6 <input type="checkbox"/> transfert d'un hôpital extérieur (or réa CHRU Lille)                  |        |

### 1 Admission

<b>Défaillance aiguë principale motivant l'admission en USC</b> <i>choix simple</i>		I __ I
1 <input type="checkbox"/> Respiratoire	7 <input type="checkbox"/> Endocrine	
2 <input type="checkbox"/> Neurologique	8 <input type="checkbox"/> Musculo-squelettique	
3 <input type="checkbox"/> Cardio-circulatoire	9 <input type="checkbox"/> Hématologique	
4 <input type="checkbox"/> Hépatique	10 <input type="checkbox"/> Multiviscérale	
5 <input type="checkbox"/> Génito-urinaire	11 <input type="checkbox"/> Indéterminée	
6 <input type="checkbox"/> Gastro-intestinale		

### Maladie Chronique

*si oui, cocher ci dessous*

<b>Défaillance(s) chroniques(s) principale(s)</b> <i>choix multiple (maximum 3)</i>		I __ I, I __ I, I __ I,
1 <input type="checkbox"/> Respiratoire	7 <input type="checkbox"/> Endocrine	
2 <input type="checkbox"/> Neurologique	8 <input type="checkbox"/> Musculo-squelettique	
3 <input type="checkbox"/> Cardio-circulatoire	9 <input type="checkbox"/> Hématologique	
4 <input type="checkbox"/> Hépatique	10 <input type="checkbox"/> Chromosomique	
5 <input type="checkbox"/> Génito-urinaire	11 <input type="checkbox"/> Indéterminée	
6 <input type="checkbox"/> Gastro-intestinale		

<b>Etiologie de cette défaillance principale</b> <i>choix simple</i>		I __ I
1 <input type="checkbox"/> Infection	6 <input type="checkbox"/> Oncologie	
2 <input type="checkbox"/> Traumatisme	7 <input type="checkbox"/> Diabète	
3 <input type="checkbox"/> Congénitale	8 <input type="checkbox"/> Allergie/Immunologie	
4 <input type="checkbox"/> Intoxication médicam.	9 <input type="checkbox"/> Mixte ou indéterminée	
5 <input type="checkbox"/> Toxicomanie		

<b>Pronostic de la maladie chronique</b> <i>choix simple</i>	I __ I (x 0 x si non applicable)
1. <input type="checkbox"/> Pas de réduction évidente de l'espérance de vie	
2. <input type="checkbox"/> Réduction de l'espérance de vie de plus de 10 ans	
3. <input type="checkbox"/> Affection habituellement létale dans l'enfance	

Score POPC avant l'admission (entre 1 et 5) I \_\_ I

### 2. Bilan du séjour et devenir :

<b>Affection caractéristique du séjour</b> <i>ch. simple</i>		I __ I
1 <input type="checkbox"/> Médicale	2 <input type="checkbox"/> Chirurgicale programmée	
	3 <input type="checkbox"/> Chirurgicale NON programmée	

Date de sortie : \_\_ / \_\_ / \_\_

Heure de sortie : \_\_ h \_\_

POPC à la sortie : I \_\_ I

Mode de sortie de l'USC : I \_\_ I

1  Vivant 2  Décédé

<b>Destination : choix simple</b>	I __ I
1 <input type="checkbox"/> Domicile	
<b>Transfert dans le même CHG :</b>	
2 <input type="checkbox"/> service de pédiatrie, soins courants	
3 <input type="checkbox"/> service de néonatalogie	
4 <input type="checkbox"/> autre service	
<b>Transfert dans un autre CHG :</b>	
5 <input type="checkbox"/> service de pédiatrie, soins courants	
6 <input type="checkbox"/> service de néonatalogie	
7 <input type="checkbox"/> autre service	
<b>Transfert au CHRU de Lille :</b>	
8 <input type="checkbox"/> service des urgences pédiatriques	
9 <input type="checkbox"/> service de spécialité péd. (cardio, chir etc)	
10 <input type="checkbox"/> service de réanimation péd CHRU Lille	
11 <input type="checkbox"/> autre service	
12 <input type="checkbox"/> Transfert dans un autre service : .....	

Si transfert en réanimation :

Mode de sortie de réa : 1  Vivant 2  Décédé I \_\_ I

0 : si pas de transfert

Score	Catégorie	Cérébral (PCPC)	POPC
1	Normal	Niveau de développement et/ou de scolarisation adaptés à l'âge.	PCPC=1 et capable d'activités normales de la vie de tous les jours
2	Déficit ou handicap mineur	Conscient, alerte, capable d'activités indépendantes appropriées pour l'âge, retard scolaire en cursus normal; déficit neurologique mineur	PCPC=2 et/ou handicap physique mineur compatible avec une vie normale et indépendante pour l'âge
3	Déficit ou handicap moyen	Conscient, capable d'activités indépendantes de la vie de tous les jours, classe d'éducation spéciale ou retard des apprentissages	PCPC=3 et/ou handicap physique modéré compatible avec les activités de la vie de tous les jours mais entraînant des performances réduites
4	Déficit ou handicap sévère	Conscient, mais dépendant des autres pour les activités de la vie de tous les jours en raison d'un déficit neurologique central	PCPC=4 et/ou handicap physique le rendant dépendant des autres pour les activités de la vie de tous les jours
5	Etat végétatif ou coma	Tout stade de coma sauf mort cérébrale. Absence d'interactivité avec son environnement	PCPC = 5

# Annexe 6 : Feuille de surveillance standardisée

N° inclusion I _ / I _ / I _ / I _		FEUILLE DE SURVEILLANCE JOURNALIERE USC		Feuille N°	
Etiquette Sigma	Heures	IDE :		IDE :	
		DATE :		Chambre ou lit : N°	
Diagnostics :	Température :				
	40°C				
	39°C				
	38°C				
	37°C				
Taille :	36°C				
Périmètre crânien :					
Allergies :					
Poids					
Hémod.	Fréquence Cardiaque (/min)				
	Pression Art. Syst/Diast (moyenne)				
	Tps de recoloration cutanée (sec)				
Respiration	SpO <sub>2</sub> (%)				
	Fréquence Respiratoire (/min)				
	Tirage <sup>1</sup> : 0 / 1 / 2 / 3				
	Apnée (Oui/Non)				
	Air ambiant (Oui/Non)				
	L=tunettes/M=masque/C=canule Débit O <sub>2</sub> (L/min) ou FIO <sub>2</sub> (%) si connue				
Neuro	Conscient/Voix/Douleur/Inconscient <sup>2</sup>				
	Pupilles réact. Symét. (Oui/Non)				
Appel du médecin ou médecin venu: O/N					
demande d'examen paracr. suppl.: O/N					
prescription d'un traitement suppl.: O/N					
transfert en REA : O/N					
Digestif	Vomissement				
	Selles / Gaz				
	Aspiration				
	Drainage				
	Diarrhée				
Rein	DIURESE				
E.V.A.	EVENDOL				
Pose K.T.	surveillance				
Aérosols					
Traitement injectable					
Traitement per os					
Soins, nutrition, voie rectale					

<sup>1</sup>Tirage respiratoire : 0 = absent  
1 = minime  
2 = modéré  
3 = sévère

<sup>2</sup>Etat de conscience : C = Conscient  
V = ne réagit qu'à la Voix  
D = ne réagit qu'à la Douleur  
I = Inconscient

## Annexe 7 : Gestion des données manquantes

**Tableau A1 : Analyse de sensibilité pour déterminer l'impact de l'imputation des données manquantes par une valeur normale comparée à l'analyse des scores complets**

	Scores complets et imputés avec une valeur normale par défaut	Scores complets
Scores	AUC (95% IC)	AUC (95% IC)
≤ 24 heures	n=4,693 observations; appel du médecin = 494 (11%)	n=4,127 observations appel du médecin = 356 (9%)
PAWS	0,87 (0,85-0,89)	0,88 (0,86-0,90)
PEWS	0,87 (0,86-0,89)	0,87 (0,85-0,90)
Bedside PEWS	0,87 (0,85-0,89)	0,85 (0,83-0,88)
≤ 36 heures	n=5,892 observations; appel du médecin (9%)	n=5,139 observations appel du médecin = 400 (8%)
PAWS	0,90 (0,88-0,91)	0,89 (0,87-0,91)
PEWS	0,89 (0,88-0,91)	0,90 (0,88-0,91)
Bedside PEWS	0,89 (0,87-0,91)	0,88 (0,85-0,90)
≤ 48 heures	n=7,196 observations appel du médecin = 629 (9%)	n= 6,284 observations appel du médecin =457 (7%)
PAWS	0,91 (0,90-0,92)	0,90 (0,88-0,92)
PEWS	0,91 (0,89-0,92)	0,91 (0,89-0,82)
Bedside PEWS	0,91(0,89-0,92)	0,89 (0,87-0,91)
≤ 6 jours	n=10,152 observations; appel du médecin = 749 (7%)	n= 8,848 observations appel du médecin =562 (6%)
PAWS	0,91 (0,90-0,92)	0,91 (0,90-0,92)
PEWS	0,91 (0,90-0,92)	0,91 (0,89-0,92)
Bedside PEWS	0,91 (0,90-0,92)	0,91 (0,89-0,92)

AUC: Aire sous la courbe ; IC: intervalle de confiance ; PAWS: Pediatric Advanced Warning Score; PEWS: Pediatric Early Warning Score; BedPEWS: Bedside Pediatric Early Warning System

## Annexe 8: Note d'information pour les parents de patients

Note d'information pour les parents des participants à une recherche observationnelle

Recherche d'un indicateur de gravité et d'activité simplifié pour les unités de surveillance continue pédiatriques : étude prospective multicentrique régionale

Organisme responsable de l'étude : CHRU de Lille

Responsable scientifique de l'étude : Pr Stéphane Leteurre

Coordonnateur de l'étude : Pr Francis Leclerc,

Service de réanimation pédiatrique, Hôpital Jeanne de Flandre, CHRU de Lille



Centre Hospitalier Régional  
Universitaire de Lille

«L'unité de surveillance continue pédiatrique du centre hospitalier de .....dans lequel est hospitalisé votre enfant participe à une étude multicentrique de la population des enfants hospitalisés dans les unités de surveillance continue de 9 hôpitaux de la région Nord-Pas de Calais. L'objectif principal de l'étude est de valider un score de gravité appréciant de façon quantitative l'état de santé des enfants. Ainsi, dans le futur, ce score pourra être utilisé comme outil d'alerte afin de détecter précocement l'amélioration ou l'aggravation de l'état de santé des enfants.

L'étude n'utilise que les données contenues dans le dossier médical de l'enfant et n'implique aucune intervention supplémentaire (aucun examen, aucun traitement supplémentaire). L'utilisation de ces données est ensuite anonyme.

Après anonymisation, les informations enregistrées feront l'objet d'un traitement informatique. Conformément aux articles 39 et suivants de la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, toute personne peut obtenir communication et, le cas échéant, rectification ou suppression des informations la concernant, en s'adressant au médecin responsable de l'étude.

Toute personne peut également s'opposer au traitement des données la concernant (ou son enfant).

Vous êtes libre de refuser de faire participer votre enfant à cette étude sans en donner la justification. Votre refus ne modifiera pas vos rapports avec les médecins et la qualité de la prise en charge de votre enfant. Vous restez libre d'arrêter la participation à tout moment.

Si vous refusez que votre enfant participe à l'étude, il suffit de l'indiquer oralement à l'un des membres de l'équipe prenant en charge votre enfant.

Le Dr....., responsable local de l'étude, est à votre disposition pour tout renseignement sur ce sujet. Vous pouvez demander à le rencontrer. Vous pourrez lui demander d'être informés des résultats globaux à l'issue de l'étude si vous le souhaitez (loi du 04 mars 2002).



## Annexe 9: Note d'information pour les enfants

Note d'information pour les enfants en âge de lecture participants à une recherche observationnelle

Recherche d'un indicateur de gravité et d'activité simplifié pour les unités de surveillance continue pédiatriques : étude prospective multicentrique régionale

Organisme responsable de l'étude : CHRU de Lille

Responsable scientifique de l'étude : Pr Stéphane Leteurtre

Coordonnateur de l'étude : Pr Francis Leclerc,

Service de réanimation pédiatrique, Hôpital Jeanne de Flandre, CHRU de Lille



Centre Hospitalier Régional  
Universitaire de Lille

L'hôpital dans lequel tu es hospitalisé participe à une recherche pour améliorer les soins donnés aux enfants. Pour cela, nous travaillons à construire un score composé de chiffres. Plus le chiffre est grand, plus l'enfant est malade et a besoin que le médecin vienne le voir.

Cette recherche se déroule aussi dans 8 autres hôpitaux de notre région.

Nous allons uniquement recopier les informations qui sont déjà dans ton dossier médical.

Ton nom n'apparaîtra nulle part. Les informations seront enregistrées puis analysées dans un ordinateur.

Tu peux refuser de participer à cette recherche sans en donner la raison. Tu seras soigné de la même façon.

Tu peux poser toutes tes questions à ce sujet au Dr....., responsable de l'étude dans le service. Tu peux demander à le rencontrer et il pourra t'expliquer plus tard les résultats de l'étude si tu le demandes.

Nous te souhaitons de vite guérir.



## Annexe 10: formulaire d'opposition à remplir par le médecin ou le soignant

### Formulaire d'opposition

Recherche d'un indicateur de gravité et d'activité simplifié pour les unités de surveillance continue pédiatriques : étude prospective multicentrique régionale

Organisme responsable de l'étude : CHRU de Lille

Responsable scientifique de l'étude : Pr Stéphane Leteurre

Coordonnateur de l'étude : Pr Francis Leclerc,

Service de réanimation pédiatrique, Hôpital Jeanne de Flandre, CHRU de Lille

Si l'un des parents ou tuteur légal de l'enfant susceptible de participer à l'étude s'y oppose, veuillez le signifier ci-dessous :

Le médecin ou le soignant .....

déclare que l'un des parents ou tuteur légal de l'enfant .....

refuse la participation à l'étude observationnelle.

Date :.....

Signature :

## **Annexe 11: Description du modèle mixte linéaire généralisé utilisé pour la validation des scores d'alerte précoce**

Variable à expliquer : critère du jugement noté  $Y$  :

$Y_{ij}=1$  ou  $0$  (critère de jugement binaire) ou  $i=$  indice individu,  $j=$  indice temps ( $j=1, \dots, n_i$ )

-Variables explicatives :

$S=$  score étudié.

$S_{i0}=$  score à l'entrée pour l'individu  $i$ .

$S_{ij}=$  score au temps  $t_{ij}$  et  $t_{ij}=$  temps  $j$  de l'individu  $i$

-Modèle mixte généralisé :

$\pi_{ij} = \Pr(Y_{ij}=1)$

$\text{Logit}(\pi_{ij}) = \beta_0 + \beta_1 S_{i0} + \beta_2 t_{ij} + \beta_3 S_{ij} + b_{0i}$

$b_{0i}$  est un effet aléatoire

$p_{ij}$  = estimation du modèle linéaire mixte généralisé.

## Annexe 12: Descriptif des données manquantes

**Tableau A2: Caractéristiques et données manquantes des trois scores**

Caractéristiques des trois scores			
Item	PAWS	PEWS	Bedside PEWS
	7	6	7
Minimum-Maximum	0-21	0-9	0-26
Items manquants pour les trois scores			
	N1/N2	N1/N2	N1/N2
Cardiovasculaire <sup>1</sup>	NA	888/924 (96,1%)	
Fréquence cardiaque	27/177 (15,2%)	NI	25/177(14,1%)
Pression artérielle systolique	NI	NI	1001/1399 (71,5%)
Temps de recoloration cutanée	938/1263 (74,3%)	NI	833/1263 (65,9%)
Respiratoire <sup>2</sup>	NI	2119/2155(98,3%)	NA
Fréquence respiratoire	275/463 (59,4%)	NI	260/463 (56,1%)
Effort respiratoire	192/496 (38,7%)	NI	115/496 (23,2%)
Saturation Oxygène <sup>3</sup>	65/243 (26,7%)	NI	62/243 (25,5%)
Comportement	64/322 (19,9%)	286/322 (88,8%)	NI
Température	422/652 (64,7%)	NI	NI
Total N1/N2 pour chaque score	1983/3616 (53,6%)	3293/3401 (96,8%)	2578/4738 (54,4%)
Total des observations pour le calcul des trois scores, n=14708			
Scores imputés	1710 (11,6%)	2916 (19,8%)	2078 (14,1%)
Scores complets	12668 (86,1%)	11756 (79,9%)	12191 (82,9%)
Scores analysés	14378 (97,8%)	14672 (99,8%)	14269 (97%)
Scores non analysés	330 (2,2%)	36 (0,2%)	439 (3%)

PAWS: Paediatric Advanced Warning Score; PEWS: Paediatric Early Warning Score; Bedside PEWS: Bedside Paediatric Early Warning System

N1/N2: nombre d'items imputés /total d'items manquants, NI: items non inclus dans le scores

<sup>1</sup> Cardiovasculaire= temps de recoloration cutanée OU fréquence cardiaque pour le PEWS; <sup>2</sup> Respiratoire: fréquence respiratoire ET effort respiratoire OU oxygène (débit/FiO2) pour le PEWS; <sup>3</sup> oxygène: room air / oxygen ; débit OU FiO2 dans le Bedside PEWS.

## Annexe 13 : Effet centre

**Tableau A3 : Comparaison des caractéristiques des patients entre les sept centres**

	Centre 1	Centre 2	Centre 3	Centre 4	Centre 5	Centre 6	Centre 7	p
Age (mois), médiane [IQR]	14 [4-56]	17[4-65]	53[18-159]	20[5-62]	92[19-181]	24[3-65]	24[3-90]	< 10 <sup>-4</sup>
Durée de séjour (jours), médiane [IQR]	1,8[1,0-3,2]	1,4[0,8-2,0]	1,2[0,8-2,0]	1,5[0,7-2,5]	1,6[0,8-2,9]	2,4[1,0-3,8]	0,9[0,7-1,9]	< 10 <sup>-4</sup>
Comorbidités, n(%)	98 (27)	74(22)	180 (38)	69 (28)	127 (47)	158 (34)	170 (36)	< 10 <sup>-4</sup>
Appel du médecin	91 (4,0)	55 (3,5)	264 (13,3)	53 (4,3)	254 (14,7)	118 (3,2)	229 (10,1)	< 10 <sup>-4</sup>
Transfert réanimation n(%)	14 (3,8)	3 (0,9)	6 (1,3)	3 (1,2)	9 (3,3)	24 (5,2)	9 (1,9)	4,10 <sup>-4</sup>
Score à admission, médiane [IQR]								
PAWS	4 [2-5]	3[2-5]	3[2-4]	2[1-4]	2[0-4]	3[2-5]	2[1-4]	< 10 <sup>-4</sup>
PEWS	2[0-3]	1[0-3]	1[0-3]	1[0-3]	0[0-2]	2[0-3]	2[1-4]	< 10 <sup>-4</sup>
Bedside PEWS	4[2-7]	3[2-6]	3[1-6]	3[1-6]	2[1-4]	4[2-8]	3[2-6]	< 10 <sup>-4</sup>

*IQR: Interquartile*, PAWS: **Paediatric Advanced Warning Score**; PEWS: Paediatric Early Warning Score; Bedside PEWS: **Bedside Paediatric Early Warning System**

**Tableau A4: Effet centre (ajustement sur le centre)**

	Avec ajustement sur le centre AUC (95%IC)	Sans ajustement sur le centre AUC (95%IC)
PAWS	0,899 (0,888-0,910)	0,914 (0,903-0,925)
PEWS	0,899 (0,887-0,910)	0,910 (0,899-0,921)
Bedside PEWS	0,897 (0,884-0,911)	0,910 (0,899-0,921)

AUC: Area Under the *receiver-operating* Curve; IC: intervalle de confiance; PAWS: **Paediatric Advanced Warning Score**; PEWS: Paediatric Early Warning Score; Bedside PEWS: **Bedside Paediatric Early Warning System**

## Annexe 14 : Données DIM

**Tableau A5: Fichier standardisé fourni aux DIM de chaque centre pour le recueil des diagnostics**

. fichier diagnosti c	N° ANONYMISAT ION	SUPPLEME NTS USC	COD E DIAG	LIBEL LE DIAG	TYPE DIAG	SUPPLEME NTS REA		
		<b>LEGENDE</b>			<b>LEGENDE</b>	<b>LEGENDE</b>		
		O = attribution du supplément			P = diagnostic principal	O = attribution du supplément		
		N = pas de supplément			R = diagnostic relié	N = pas de supplément		
	1/001							
	1/002							

**Tableau A6: Fichier standardisé fourni aux DIM de chaque centre pour le recueil des Actes**

fichier actes	N° ANONYMISATION	CODE ACTE	LIBELLE ACTE	Dates Acte
	1/001			
	1/002			