

# **PaO<sub>2</sub>/FiO<sub>2</sub> ratio derived from the SpO<sub>2</sub>/FiO<sub>2</sub> ratio to improve mortality prediction using the Paediatric Index of Mortality-3 score in transported intensive care admissions**

Samiran Ray MBBChir MA<sup>1,2</sup>, Libby Rogers MMath<sup>3</sup>, Christina Pagel PhD<sup>2,3</sup>, Sainath Raman MBBS<sup>1</sup>, Mark J Peters MBChB PhD<sup>1,2</sup>, Padmanabhan Ramnarayan MBBS MD FFICM<sup>1,2</sup>

<sup>1</sup>Respiratory Critical Care and Anaesthesia Unit, UCL Great Ormond Street Institute of Child Health, London

<sup>2</sup>Children's Acute Transport Service, Great Ormond Street Hospital, London UK

<sup>3</sup>UCL Clinical Operational Research Unit, London UK

## **Corresponding Author:**

Dr Samiran Ray

Children's Acute Transport Service,

Great Ormond Street Hospital,

London WC1N 3JH

Tel: (044) 2074059200

Fax: (044) 2078297830

e-mail: Samiran.ray@gosh.nhs.uk

**Key words:** oximetry; blood gas monitoring; child; critical care; prognosis; transport

## **Abstract**

Objective: To derive a relationship between the  $\text{SpO}_2/\text{FiO}_2$  ratio (SF) and  $\text{PaO}_2/\text{FiO}_2$  ratio (PF) across the entire range of  $\text{SpO}_2$  values (0-100%) and to evaluate whether mortality prediction using the Paediatric Index of Mortality (PIM-3) can be improved by the use of PF values derived from SF.

Design: Retrospective analysis of prospectively collected data.

Setting: A regional paediatric intensive care (PICU) transport service.

Patients: Children transported to a PICU.

Interventions: None

Measurements and Main Results: The relationship between SF and PF across the entire range of  $\text{SpO}_2$  values was first studied using several mathematical models in a derivation cohort ( $n=1235$ ) and then validated in a separate cohort ( $n=306$ ). The best SF to PF relationship was chosen according to the ability to detect respiratory failure ( $\text{PF} \leq 200$ ). The discrimination of the original PIM-3 score and a derived PIM-3 score (where SF-derived PF values were used in place of missing PF values) were compared in a different cohort ( $n=1205$ ). The best SF-PF relationship in 1703 SF/PF data pairs was a linear regression equation of  $\ln[\text{PF}]$  regressed on  $\ln[\text{SF}]$ . This equation identified children with a  $\text{PF} \leq 200$  with a specificity of 73% and sensitivity of 61% in children with  $\text{SpO}_2 < 97\%$  (92% and 33% respectively when  $\text{SpO}_2 \geq 97\%$ ) in the validation cohort. PF derived from SF (dPF) was better at predicting PICU mortality (AUROC 0.64, 95% CI 0.55-0.73) compared to the original PF (AUROC 0.54, 95% CI 0.49-0.59,  $p=0.02$ ). However there was no difference in the original and derived PIM-3 scores and their discriminatory ability for mortality.

Conclusion:  $\text{SpO}_2$  based metrics perform no worse than arterial blood gas based metrics in mortality prediction models. Future PIM score versions may be improved by the inclusion of

risk factors based on oxygen saturation values, especially in settings where PaO<sub>2</sub> values are missing in a significant proportion of cases.

## **Introduction**

Respiratory failure is the commonest cause for unplanned intensive care admission in children in the United Kingdom (UK) (1). The  $\text{PaO}_2/\text{FiO}_2$  (PF) ratio has traditionally been used as a marker of lung injury – it quantifies the capacity for gas exchange at the alveolar surface. The PF ratio has a strong relationship with severity of illness and mortality: it is therefore included as a variable in most adult (2, 3) and paediatric (4-6) severity of illness scores.

In recent years, however, the ability to monitor haemoglobin oxygen saturations ( $\text{SpO}_2$ ) non-invasively and continuously using pulse oximetry has led to a decline in the use of arterial blood gas analysis to monitor oxygenation status (7). In recognition of this, several studies have investigated the relationship between PF and the ratio of  $\text{SpO}_2$  to  $\text{FiO}_2$  (SF) (8, 9) in paediatric intensive care unit (PICU) patients. Thresholds based on SF have recently been proposed as an alternative to PF thresholds in the definition of paediatric acute respiratory distress syndrome (pARDS) (10). However, these validation studies excluded  $\text{SpO}_2$  values greater than 97%, given that the oxygen dissociation curve plateaus beyond this point. Since a large proportion of PICU patients have  $\text{SpO}_2$  values greater than 97% (57% of children in one study assessing the use of  $\text{SpO}_2$  in the PIM-2 score) (11), SF-derived PF values cannot currently be calculated for the majority of PICU patients.

Paediatric severity of illness scores such as the Paediatric Index of Mortality (PIM) that rely on PF for score calculation use different techniques to account for PF values that are missing – PIM-2 assumes a PF value of 0, while PIM-3 assumes a ‘normal’ PF value of 0.23 (based on a  $\text{PaO}_2$  value of 91 mm Hg or 12 kPa when breathing room air) (4, 12). However PF values

are rarely missing at random in this population – children with available PF values are likely to have a higher illness severity, as sicker children are more likely to have arterial cannulation. This introduces potential bias in the calculation of the original PIM-3, leading to over or under-estimation of mortality. Leteurtre et al tested the use of SF-derived PF estimates to replace missing PF values during the calculation of the PIM-2 score in their PICU population: this modified PIM-2 score performed at least as well as the original PIM-2 score in predicting mortality (11). The aims of our study were to: (a) derive a relationship between SF and PF across *all* SpO<sub>2</sub> values, including over 97%, (b) validate it in a separate population, and (c) test the predictive value of PIM-3 scores calculated using SF-derived PF estimates, in a population of children transported into PICU, where the number of missing PF values is high.

## Methods

We conducted a retrospective observational study using data from all children transported to PICU by a regional transport team (Children's Acute Transport Service, London, UK) between April 2011 and October 2014. This service retrieves children requiring intensive care from over 30 referral hospitals into 4 regional general PICUs and 2 cardiac PICUs. Clinical practice relating to mechanical ventilation during transport was generally consistent with the practices of the regional PICUs, including use of lung protective strategies such as low tidal volume ventilation, use of PEEP and permissive hypercarbia.

The study was registered with the local institutional review board. The need for patient consent was waived by the Yorkshire and Humber Research Ethics Committee (IRAS 191836). Collection of personally identifiable data by the Paediatric Intensive Care Audit

Network (PICANet) was approved by the NHS Health Research Authority Confidentiality Advisory Group and ethics approval granted by the Trent Medical Research Ethics Committee, ref. 05/MRE04/17.

#### Derivation dataset:

The derivation dataset consisted of PaO<sub>2</sub>, FiO<sub>2</sub> and concurrent SpO<sub>2</sub> values recorded in the case-notes of children (0-16 years of age), who had arterial blood gas analysis performed during transport between April 2011 and October 2013 (30 months). We used PaO<sub>2</sub>, FiO<sub>2</sub> and concurrent SpO<sub>2</sub> values from all blood gases performed during transport to derive a relationship between SF and PF; we did not exclude SpO<sub>2</sub> values >=97%.

#### Validation dataset:

The validation dataset consisted of PaO<sub>2</sub>, FiO<sub>2</sub> and concurrent SpO<sub>2</sub> values recorded in the case notes of children (0-16 years of age) who had arterial blood gas analysis during transport between October 2013 and April 2014 (6 months). The best relationship derived in the derivation cohort was validated in this cohort.

#### Testing dataset:

The test dataset consisted of PaO<sub>2</sub> (if available), FiO<sub>2</sub> and concurrent SpO<sub>2</sub> values recorded in the first hour of contact with the PIC transport team (as required by PIM (4)) for all children transported to PICU between October 2013 and 2014 (12 months). For these patients, survival status at discharge from PICU, the original PIM-3 scores and the PaO<sub>2</sub> and FiO<sub>2</sub> values (if available) used to calculate the original PIM-3 scores, were obtained from PICANet ([www.picanet.org.uk](http://www.picanet.org.uk)).

### Statistical analysis:

Derivation and validation of the relationship between SF and PF: SF and PF ratios were calculated by dividing  $\text{SpO}_2$  by  $\text{FiO}_2$ , and  $\text{PaO}_2$  (in mmHg) by  $\text{FiO}_2$ , respectively. We investigated the best relationship between SF and PF among the following multi-level regression models (patients as levels): (a) linear regression of PF on SF , (b) piecewise linear regression of PF on SF, allowing different slopes for  $\text{SpO}_2 < 94$ ,  $\text{SpO}_2 94-97$  and  $\text{SpO}_2 \geq 97$  to take into account the shape of the oxygen dissociation curve, (c) fractional polynomial regression of PF and SF, regressing PF on different combination of powers of SF, again to take into account the non-linear relationship of the oxygen dissociation curve (d)-(f) methods a-c were repeated following natural log transformation of SF and PF. The best model was chosen based on sensitivity analysis for the ability to detect respiratory failure with a  $\text{PF} \leq 200$ , both with  $\text{SpO}_2 < 97$  and  $\text{SpO}_2 \geq 97$ , since identification of respiratory failure in the latter group is the determinant of mortality risk. This model was validated in the validation cohort, on the basis of the sensitivity and specificity for detecting  $\text{PF} \leq 200$  mm Hg.

dPIM3 testing: We calculated a derived PF (dPF) value for children in the test cohort using the SF-PF relationship derived and validated in the previous steps. We used the dPF value to calculate a derived PIM-3 score (dPIM-3). If either  $\text{SpO}_2$  or  $\text{FiO}_2$  values were missing the default value of 0.23 was used for  $100/\text{PF}$  (4). Receiver operating characteristic (ROC) curves were used to test the ability of dPF and PF, and PIM-3 and dPIM-3 values to predict PICU mortality. Areas under the ROC curves were compared using DeLong's method. Calibration was tested across deciles of risk using the Hosmer-Lemeshow Goodness of Fit

test. All analysis was carried out using Microsoft Excel (Microsoft Corporation, WA, USA), Stata v14.1 (StataCorp LP, TX, USA) and R ([www.cran.r-project.org](http://www.cran.r-project.org)).

## Results

### Derivation of the SF-PF relationship

A total of 2823 were transported in the 30 months period; 1235 children had arterial blood gas analysis. From these 1235 children, 1703 concurrently recorded  $\text{PaO}_2$ ,  $\text{SpO}_2$  and  $\text{FiO}_2$  values were available. The characteristics of this derivation cohort are presented in Table 1.

Figure 1 shows the relationship between SF and PF, before and after log transformation. The best relationship between PF and SF was obtained from linear regression following log transformation. The derived equation was

$$\ln(\text{PF}) = 1.086 \times \ln(\text{SF}) + 0.103$$

$$\text{Alternatively, } \text{PF} = 1.109 \times (\text{SF}^{1.086})$$

This equation had 70% sensitivity and 73% specificity in identifying children with  $\text{PF} \leq 200$  and  $\text{SpO}_2 < 97\%$ ; and 49% sensitivity and 76% specificity in identifying children with  $\text{PF} \leq 200$  and  $\text{SpO}_2 \geq 97\%$ . While other equations, such as those derived using piecewise linear fit, were more sensitive at identifying children with a  $\text{PF} \leq 200$  for  $\text{SpO}_2 \leq 97\%$ , they had poor specificity. Also, none of the other equations performed well when  $\text{SpO}_2 \geq 97\%$ , with 0% sensitivity (i.e. no children with  $\text{SpO}_2 \geq 97\%$  and  $\text{PF} \leq 200$  were identified). The high specificity suggests a high rule-out value – not all children with  $\text{dPF} \leq 200$  have respiratory failure but if they have  $\text{dPF} > 200$  they are unlikely to be have respiratory failure.

### Validation of the SF-PF relationship

A total of 935 children were transported in the 6 month period (October 2013-April 2014); 306 had arterial blood gas analysis. The derived equation was tested on 425 concurrently recorded  $\text{PaO}_2$ ,  $\text{SpO}_2$  and  $\text{FiO}_2$  values from these 306 patients (Table 1). The equation identified children with  $\text{PF} \leq 200$  with 61% sensitivity and 73% specificity in children with  $\text{SpO}_2 < 97\%$ ; and 33% sensitivity and 92% specificity in children with  $\text{SpO}_2 \geq 97\%$ . This is similar in children with respiratory disease within this cohort ( $\text{SpO}_2 < 97\%$ : sensitivity 58%, specificity 75%;  $\text{SpO}_2 \geq 97\%$ : sensitivity 30%, specificity 93%).

### Testing mortality prediction using the original and derived PIM3 scores

Outcome data were available through PICANET for 897 children out of 966 children transported to PICU (69 cases could not be matched to a PICU admission by PICANET). Out of 897 children, 46 died prior to PICU discharge (5.1%). PF data were available for 113/897 (12.6%); in contrast, SF data were available for 819/897 (91.3%) children (Table 2).

Using the derived equation, we calculated dPF in the validation cohort. Figure 2 shows the ROC curves for (a) the component of the PIM-3 coefficient using PF and dPF ( $0.4214 * 100 / \text{PF}$  or  $0.4214 * 0.23$  if PF missing) and (b) PIM-3 and dPIM-3 values. While there was a significant difference in the area under the ROC curve between the PF and dPF ( $p\text{-value}=0.02$ ) with the latter performing better, there was no significant difference in PIM- and dPIM-3 values ( $p\text{-value}=0.31$ ). When calibration was tested using the Hosmer-Lemeshow test across deciles of risk, there was no difference between PIM-3 and dPIM3 scores (data tables in ESM).

### Discussion:

In this retrospective observational study, we derived and validated a relationship between SF and PF using data from a cohort of patients being transported to PICU. Unlike previously reported relationships, we included all values of SpO<sub>2</sub>, including >=97%. We found that the SF-derived PF component of PIM-3 as a single variable predicted mortality better than the original PF. However, when the SF-derived PF was incorporated within PIM-3, both the original and the derived PIM-3 scores had similar discrimination.

Our derived relationship between SF and PF is novel in that it includes children with SpO<sub>2</sub>>=97%. Previous derivations in children have excluded children with SpO<sub>2</sub>>=97% to represent only the linear part of the oxygen dissociation curve. However a high proportion of children in the intensive care have SpO<sub>2</sub>>=97% (59% in one study (11)). In our derivation and validation cohort, 75% of children had SpO<sub>2</sub>>=97%; similarly 67% in the testing cohort. This observation holds true even in children with significant lung disease - we recently reported that nearly 50% of children with a PF<=300 mm Hg had SpO<sub>2</sub> values>=97% (14). For PIM, which applies to a general population, this is important. Although the sensitivity in children with SpO<sub>2</sub>>=97% is poor (49% in derivation and 33% in validation cohorts) compared to the sensitivity of 68% reported by Khemani et al (8) when values of SpO<sub>2</sub>>=97% were excluded, many of these children may not have arterial gas analysis. Currently PIM-3 would treat children with missing values as having normal pulmonary gas exchange which has zero sensitivity.

The fact that the PIM-3 value does not change, even when there is a difference in the PF and dPF values, may be due to the relatively low weight that is placed on the PF component in the PIM equation. It is possible that if SF ratios were used prospectively to derive the PIM

equation with better data completeness, then PIM would have greater sensitivity for hypoxia at PICU admission.

The area under the ROC curve for predicting PICU mortality was 0.64 for dPF as used in PIM-3. This suggests that although modest, the SF ratio at admission may have some predictive value in predicting PICU outcome: if this was not the case, then there would be little value in using SF derived PIM scores. We have previously shown this in a larger cohort of children (13). The poor performance of PF in predicting mortality is likely due to the large number of missing values: only 12.6% of our cohort of transported children had PF values (compared to 77.3% in a population admitted post cardiac surgery from a single centre over the same period). In contrast SF values are available in 91% of the transported children. Therefore the use of dPIM-3 is likely to be most valuable in such a cohort. The use of SF will preclude the prediction of any mortality risk from hyperoxia, although PIM-3 currently does not do so either.

Our study has some limitations. We used data from children undergoing emergency admission following transport from a local hospital. PIM-3 applies to all children admitted to PICU, therefore the use of SpO<sub>2</sub> based equations will need testing across the population. Oxygen saturation readings are less accurate at the <80-85% range (15) and dependent on the manufacturer of the probe. Although we currently use the Massimo SET with LNCS sensor (Massimo, CA, USA), the first SpO<sub>2</sub> values in children transported from other hospitals may have been from other monitors. While this may reduce reliability, it makes our derivation more generalizable. In our derivation set we used multiple concurrent PF and SF pairs for some patients i.e. not restricted to the first hour of contact with the intensive

care team. In reality the first point of contact with the intensive care team – the time at which PIM is calculated – is variable. Our median retrieval time is under two hours, with almost all retrievals lasting less than four hours. Therefore all the blood gases in the derivation set would have been within four hours of contact with the intensive care team.

Pulse oximetry has become the more widely used indicator of oxygenation compared to arterial blood gas analysis in the general PICU population (16). We propose this should be reflected in predictive and severity scoring. Although PIM-3 acknowledges the impact of missing PF data better than previous iterations, the use of SF data is the most logical alternative: it performs no worse than PF, and with new calibration, may well improve severity of illness scoring. With only 12% of transported patients having PF values and the tendency towards less invasive monitoring in PICU patients overall, the use of SF in PIM scoring is likely to be of better discriminatory value.

**Acknowledgements:** This work was undertaken at UCL GOS Institute of Child Health, which received a proportion of funding from the Department of Health's NIHR Biomedical Research Centre's funding scheme.

The PICANet Audit is commissioned by the Healthcare Quality Improvement Partnership (HQIP) as part of the National Clinical Audit Programme (NCA). HQIP is led by a consortium of the Academy of Medical Royal Colleges, the Royal College of Nursing and National Voices. Its aim is to promote quality improvement, and in particular to increase the impact that clinical audit has on healthcare quality in England and Wales. HQIP holds the contract to manage and develop the NCA Programme, comprising more than 30 clinical audits that cover care provided to people with a wide range of medical, surgical and mental health

conditions. The PICANet Audit is funded by NHS England, NHS Wales, NHS Lothian/National Service Division NHS Scotland, the Royal Belfast Hospital for Sick Children, The National Office of Clinical Audit (NOCA), Republic of Ireland and HCA Healthcare.

### **References:**

1. Universities of Leeds and Leicester. Paediatric Intensive Care Audit Network (PICANet) 2015 Annual Report. Available at:  
[http://www.picanet.org.uk/Audit/Annual-Reporting/PICANet\\_2015\\_Annual\\_Report\\_Summary.pdf](http://www.picanet.org.uk/Audit/Annual-Reporting/PICANet_2015_Annual_Report_Summary.pdf). Accessed May 2016
2. Zimmerman JE, Kramer AA, McNair DS et al. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Crit Care Med.* 2006 May;34(5):1297-310.
3. Vincent JL, Moreno R, Takala J et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med.* 1996 Jul;22(7):707-10.
4. Straney L, Clements A, Parslow RC et al.; ANZICS Paediatric Study Group and the Paediatric Intensive Care Audit Network. Paediatric index of mortality 3: an updated model for predicting mortality in pediatric intensive care\*. *Pediatr Crit Care Med.* 2013 Sep;14(7):673-81.
5. Leteurtre S, Duhamel A, Salleron J et al.; Groupe Francophone de Réanimation et d'Urgences Pédiatriques (GFRUP). PELOD-2: an update of the PEdiatric logistic organ dysfunction score. *Crit Care Med.* 2013 Jul;41(7):1761-73.

6. Pollack MM, Patel KM, Ruttmann UE. PRISM III: an updated Pediatric Risk of Mortality score. *Crit Care Med.* 1996 May;24(5):743-52.
7. Khemani RG, Rubin S, Belani S et al. Pulse oximetry vs. PaO<sub>2</sub> metrics in mechanically ventilated children: Berlin definition of ARDS and mortality risk. *Intensive Care Med.* 2015 Jan;41(1):94-102.
8. Khemani RG, Thomas NJ, Venkatachalam V et al; Pediatric Acute Lung Injury and Sepsis Network Investigators (PALISI). Comparison of SpO<sub>2</sub> to PaO<sub>2</sub> based markers of lung disease severity for children with acute lung injury. *Crit Care Med.* 2012 Apr;40(4):1309-16.
9. Lobete C, Medina A, Rey C et al Correlation of oxygen saturation as measured by pulse oximetry/fraction of inspired oxygen ratio with Pao<sub>2</sub>/fraction of inspired oxygen ratio in a heterogeneous sample of critically ill children. *J Crit Care.* 2013 Aug;28(4):538.e1-7.
10. Pediatric Acute Lung Injury Consensus Conference Group. Pediatric acute respiratory distress syndrome: consensus recommendations from the Pediatric Acute Lung Injury Consensus Conference. *Pediatr Crit Care Med.* 2015 Jun;16(5):428-39.
11. Leteurtre S, Dupré M, Dorkenoo A et al Assessment of the Pediatric Index of Mortality 2 with the Pao<sub>2</sub>/Fio<sub>2</sub> ratio derived from the Spo<sub>2</sub>/Fio<sub>2</sub> ratio: a prospective pilot study in a French pediatric intensive care unit. *Pediatr Crit Care Med.* 2011 Jul;12(4):e184-6.
12. 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 Feb;29(2):278-85.

13. Raman S, Ray S, Prince N et al. Hyperoxia, hypoxia and mortality in critically ill children: A cohort study and systematic review. *Pediatr Crit Care Med.* 2016 Oct;17(10):e444-e450
14. S Ray, L Rogers, S Raman, MJ Peters on behalf of the Oxy-PICU investigators. Liberal oxygenation in paediatric intensive care: retrospective analysis of high resolution SpO<sub>2</sub> data. *Intensive Care Med.* 2016 Oct 28. [Epub ahead of print]
15. Harris BU, Char DS, Feinstein JA et al. Accuracy of Pulse Oximeters Intended for Hypoxemic Pediatric Patients. *Pediatr Crit Care Med.* 2016 Apr;17(4):315-20.
16. Khemani RG, Markovitz BP, Curley MA. Characteristics of children intubated and mechanically ventilated in 16 PICUs. *Chest.* 2009 Sep;136(3):765-71.

**Figure 1:** The relationship between  $\text{PaO}_2/\text{FiO}_2$  and  $\text{SpO}_2/\text{FiO}_2$  before (above) and after (below) log transformation.  $\text{PaO}_2$  is measured in mm Hg. The line of best fit following log transformation is given by the following equation:

$$\ln(\text{PF}) = \{1.271 \times \ln(\text{SF})\} - 0.864 \text{ (black line)}$$

However following multi-level linear regression analysis to account for multiple pairs from the same patient, the equation is modified to:

$$\ln(\text{PF}) = \{1.086 \times \ln(\text{SF})\} + 0.103 \text{ (gray dotted line)}$$

**Figure 2:** Receiver operating characteristic (ROC) curves for (a) the derived and original PF component and (b) the derived and original PIM-3 scores. The area under the ROC curve was significantly different for the PF component (DeLong's p-value=0.02, denoted \*): this is likely to be due to the way missing values are accounted for in PIM-3. However the area under the ROC curve was not different for the original and derived PIM-3 scores (DeLong's p-value=0.31): this is likely a reflection of the relatively small coefficient assigned to PF in the PIM-3 equation.

**Table 1**

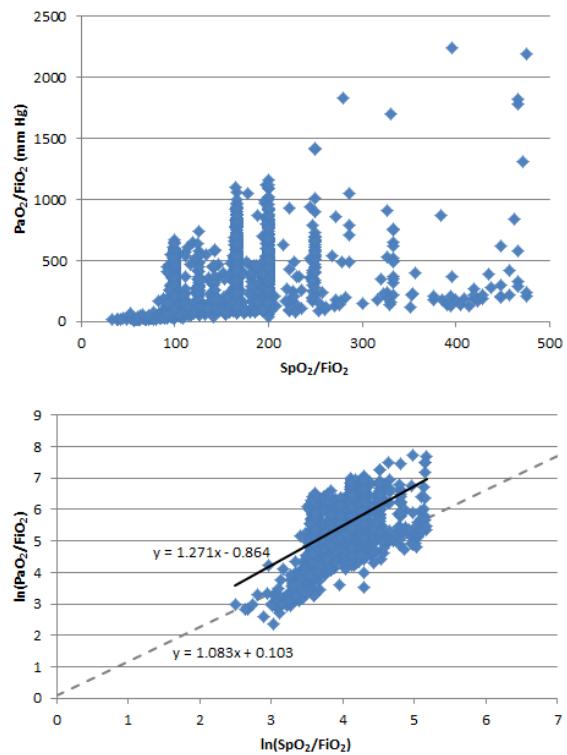
	<b>Derivation cohort</b> (April 2011 – October 2013)	<b>Validation cohort</b> (October 2013 – April 2014)
Total number of children	1235	306
Age, n (%)		
0-1 month	272 (22.0)	66 (21.6)
1-6 months	166 (13.4)	65 (21.2)
6-12 months	124 (10.0)	33 (10.8)
1-2 years	163 (13.2)	29 (9.5)
2-5 years	212 (17.2)	45 (14.7)
>5 years	298 (24.1)	68 (22.2)
Diagnostic category, n (%)		
Respiratory	486 (39.4)	152 (49.7)
Cardiovascular	146 (11.8)	40 (13.1)
Neurology	245 (19.8)	42 (13.7)
Infection/auto-immune	116 (9.4)	30 (9.8)
Other	242 (19.6)	42 (13.7)
Mechanical ventilation, n (%)	1235 (100)	306 (100)
Haemoglobin concentration (g/L), median (IQR)	109 (95-126)	108 (92-125)
Temperature (°Centigrade), median (IQR)	36.6 (35.7-37.3)	36.8 (36-37.5)
Number of concurrent PaO <sub>2</sub> and SpO <sub>2</sub> values	1703	425
PaO <sub>2</sub> :FiO <sub>2</sub> ratio (%)		
<=200 mm Hg	676 (39.7)	181 (42.6)
>200 mm Hg	1027 (60.3)	244 (57.4)
SpO <sub>2</sub> (%)		
<97%	418 (24.5)	108 (25.4)
≥97%	1285 (75.5)	317 (74.6)

**Table 1:** Baseline characteristics of children in the derivation and validation cohorts. The derivation cohort included 1703 concurrent PaO<sub>2</sub> and SpO<sub>2</sub> samples from 1235 children over 30 months. The validation cohort included 425 concurrent PaO<sub>2</sub> and SpO<sub>2</sub> samples from 306 children over 6 months. Significantly more children in the validation cohort had respiratory illnesses ( $p=0.015$  on chi-square test), which is likely as these children were transported over winter.

**Table 2**

	dPIM testing cohort (October 2013 – October 2014)
Total number of children	897
Age, n (%)	
0-1 month	201 (22.4)
1-6 months	171 (19.1)
6-12 months	100 (11.1)
1-2 years	103 (11.5)
2-5 years	139 (15.5)
>5 years	183 (20.4)
Diagnostic category, n (%)	
Respiratory	336 (37.5)
Cardiovascular	163 (18.1)
Neurology	169 (18.8)
Infection/auto-immune	79 (8.8)
Other	150 (16.7)
Mechanical ventilation, n (%)	737 (82.2)
SpO2	
<97%	274 (30.5)
>=97%	603 (67.2)
Missing SpO2 in first hour	20 (2.2)

**Table 2:** Baseline characteristics of dPIM testing cohort. The testing cohort included 897/966 children transported to PICU who had available outcome data.



**Figure 1:** The relationship between  $\text{PaO}_2/\text{FiO}_2$  and  $\text{SpO}_2/\text{FiO}_2$  before (above) and after (below) log transformation.

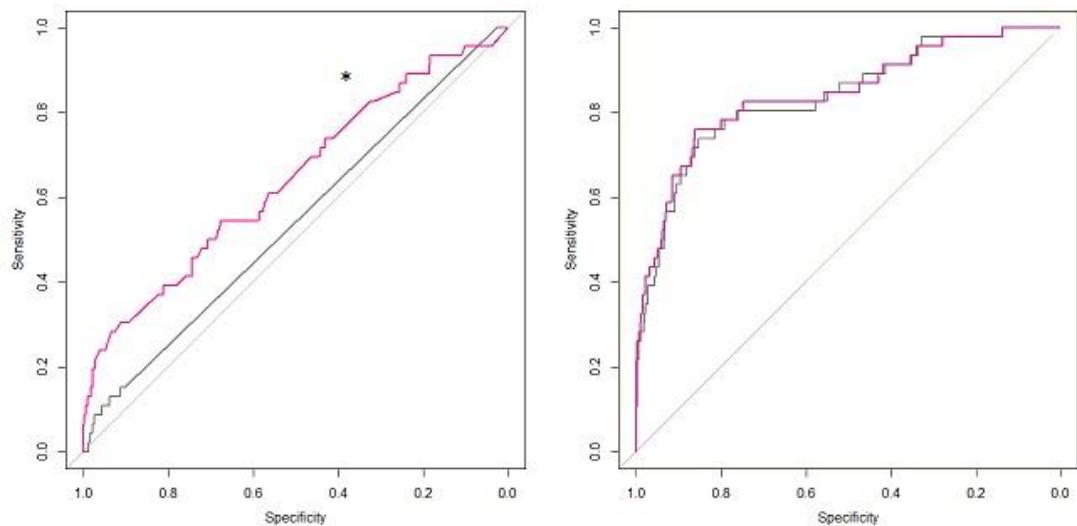
$\text{PaO}_2$  is measured in mm Hg. The line of best fit following log transformation is given by the

following equation:

$$\ln(\text{PF}) = \{1.271 \times \ln(\text{SF})\} - 0.864 \quad (\text{black line})$$

However following multi-level linear regression analysis to account for multiple pairs from the same patient, the equation is modified to:

$$\ln(\text{PF}) = \{1.086 \times \ln(\text{SF})\} + 0.103 \quad (\text{gray dotted line})$$



**Figure 2: Receiver operating characteristic (ROC) curves for (a) the derived and original PF component and (b) the derived and original PIM-3 scores. The area under the ROC curve was significantly different for the PF component (DeLong's p-value=0.02, denoted \*): this is likely to be due to the way missing values are accounted for in PIM-3. However the area under the ROC curve was not different for the original and derived PIM-3 scores (DeLong's p-value=0.31): this is likely a reflection of the relatively small coefficient assigned to PF in the PIM-3 equation.**