OPEN

Estimated Pao_2: A Continuous and Noninvasive Method to Estimate Pao_2 and Oxygenation Index

Michaël Sauthier, MD, MBI[1–3]
Gaurav Tuli, PhD[4]
Philippe A. Jouvet, MD, PhD, MBA[2]
John S. Brownstein, PhD[4]
Adrienne G. Randolph, MD, MSc[1,3]

BACKGROUND: Pao_2 is the gold standard to assess acute hypoxic respiratory failure, but it is only routinely available by intermittent spot checks, precluding any automatic continuous analysis for bedside tools.

OBJECTIVE: To validate a continuous and noninvasive method to estimate hypoxemia severity for all Spo_2 values.

DERIVATION COHORT: All patients who had an arterial blood gas and simultaneous continuous noninvasive monitoring from 2011 to 2019 at Boston Children's Hospital (Boston, MA) PICU.

VALIDATION COHORT: External cohort at Sainte-Justine Hospital PICU (Montreal, QC, Canada) from 2017 to 2020.

PREDICTION MODEL: We estimated the Pao_2 using three kinds of neural networks and an empirically optimized mathematical model derived from known physiologic equations.

RESULTS: We included 52,879 Pao_2 (3,252 patients) in the derivation dataset and 12,047 Pao_2 (926 patients) in the validation dataset. The mean function on the last minute before the arterial blood gas had the lowest bias (bias −0.1% validation cohort). A difference greater than or equal to 3% between pulse rate and electrical heart rate decreased the intraclass correlation coefficients (0.75 vs 0.44; $p <$ 0.001) implying measurement noise. Our estimated Pao_2 equation had the highest intraclass correlation coefficient (0.38; 95% CI, 0.36–0.39; validation cohort) and outperformed neural networks and existing equations. Using the estimated Pao_2 to estimate the oxygenation index showed a significantly better hypoxemia classification (kappa) than oxygenation saturation index for both Spo_2 less than or equal to 97% (0.79 vs 0.60; $p < 0.001$) and Spo_2 greater than 97% (0.58 vs 0.52; $p < 0.001$).

CONCLUSION: The estimated Pao_2 using pulse rate and electrical heart rate Spo_2 validation allows a continuous and noninvasive estimation of the oxygenation index that is valid for Spo_2 less than or equal to 97% and for Spo_2 greater than 97%. Display of continuous analysis of estimated Pao_2 and estimated oxygenation index may provide decision support to assist with hypoxemia diagnosis and oxygen titration in critically ill patients.

KEY WORDS: automatic data processing; clinical decision support systems; critical care; machine learning; oximetry

Acute hypoxic respiratory failure is a common reason of admission in PICUs and is associated with a high mortality and morbidity (1–3). The gold standard to assess hypoxemia severity in pediatric requires an arterial blood gas (ABG) to measure the oxygen partial pressure (Pao_2) and to calculate the oxygenation index (OI = [$Fio_2 \times 100 \times$ mean airway pressure]/Pao_2) (4). Although Pao_2 is the reference, it is an invasive method available only for selected patients. Furthermore, only intermittent spot checks are routinely possible, precluding accurate continuous evaluation. Oxygen is most a rapidly and widely changing analyte. Clinically, this may lead to periods of undetected hypoxemia (5).

DOI: 10.1097/CCE.0000000000000546

Pulse oximetry (Spo_2) is the most common continuous surrogate, but the relationship among Spo_2, the Sao_2, and Pao_2 is complex, especially when Spo_2 is greater than 97% (6, 7). Therefore, pediatric and adult acute respiratory distress syndrome (ARDS) guidelines recommend to use only Spo_2 less than or equal to 97% to assess hypoxemia severity (4, 8). This is an important limitation for a continuous estimation in the ICU as a majority of patients have Spo_2 above 95% (3, 9–13).

First, equations describing oxyhemoglobin dissociation curve had important limitations (14–16). They cannot estimate the Pao_2 when Sao_2 is close to 100% (asymptotic relationship) and are significantly biased for high Sao_2 values (17). Gadrey et al (17) modified equation by Hill (15) to address that limitation and showed improved performance in hospitalized non-ICU patients. Furthermore, all existing equations used a value of P_{50} (Pao_2 value when Sao_2 = 50%) that was measured on a few healthy adults (15), but this parameter is known to be different in a critically ill population (18). Brockway showed that Pao_2 can be estimated from Spo_2 in neonates (19). Other equations estimating the Pao_2/Fio_2 ratio from the Spo_2/Fio_2 ratio have been developed (20–24). Although Pao_2 can be mathematically solved, fit was generally poor (17, 20, 25).

Only few studies have used automatic continuously collected streams of data (6). This rich data source is helpful to precisely pair ABGs with other continuous monitoring data. Because no human is involved in the process, they require accurate automatic detection mechanisms to exclude potentially erroneous data (26). Different time windows and aggregation functions have been described in the literature from –60 minutes to the near-exact time (6, 17, 23, 27), but no comparative methods have been published in the literature to our knowledge.

This study aimed to validate a method to continuously estimate the hypoxemia severity, valid for all Spo_2 values, by estimating the Pao_2 using noninvasive monitoring information. Such a method would facilitate early detection and monitoring of hypoxemia.

MATERIAL AND METHODS

We collected two independent datasets at Boston Children's Hospital (BCH, Boston, MA) ICU (excluding the cardiac unit) and at Sainte-Justine Hospital (SJH, Montreal, QC, Canada) PICU. Both are academic referral centers that treat a broad range of highly specialized and common pediatric critical conditions, and they are approximately at the same altitude (< 50 m above sea level). We used the BCH cohort to develop new models and the SJH cohort to validate them independently. We included any patients admitted in these PICUs who had an ABG with concomitant continuous Spo_2 measures. ABG were excluded if more than 20% of the Spo_2 were missing before the ABG. This study was approved by both BCH (reference IRB-P00021911) and SJH (reference 2018-1587) institutional review boards and waived the need for informed consent. We followed the 2020 standards for prediction models in critical care (28) and the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis guidelines for development and validation of predictive models (29).

Data Collection

Data were extracted from secured servers located in each hospital. At BCH, data were retrieved at a 5-second frequency from the T3 database (Etiometry, Boston, MA) from October 2011 to December 2019. At SJH, data were collected from January 2017 to January 2020 and retrieved at 1-second rate from a software developed in-house (30). We collected Spo_2, pulse rate, and heart rate from –20 to +10 minutes around the ABG. SJH data were resampled (closest 5-s time) to have a simulated 5-second frequency. In both centers, ABGs were analyzed using Radiometer ABL90 FLEX (Radiometer Medical ApS, Copenhagen, Denmark). In both centers, Spo_2 was measured using Masimo (Masimo, Irvine, CA) probes (RD SET Neo and RD SET Inf). Body locations of the probe were not available as continuous data. Once extracted, data were stored and managed on a secured local PostgreSQL 12.1 server (The PostgreSQL Global Development Group) and were analyzed using Python 3.7.7 (Python Software Foundation, Wilmington, DE) and R 3.6.3 (R Core Team, Vienna, Austria) scripts. Neural networks models were built with Keras 2.3.0 on TensorFlow 2.1.0 (Google, Mountain View, CA). In the models, we only used Spo_2, heart rate, and pulse rate.

Statistical Analyses and Aggregation Function

Categorical data were described as count and percentage, and continuous data as median and interquartile range. CIs and *p* values were calculated using 10,000 bootstrap repetitions (31). We used the intraclass correlation

coefficient (ICC) as a global metrics for agreement, accuracy, and precision (32). We used the two-way mixed effects ICC, also known as "ICC" (3,1) in the Shrout and Fleiss (33) classification. The accuracy was estimated using the bias defined as the mean of the differences (mean biased error or fixed bias) (34). Precision was estimated with the mean absolute error (MAE) and the limits of agreements (95% CI on the differences) (34). Agreement was shown with Bland-Altman plots grouped by a range of Spo_2. We also calculated the proportional bias, defined by the slope on the Bland-Altman plot. Discrimination was estimated using the area under the receiving operating characteristic curve (AUROC) for mild (OI ≥ 4) and severe (OI ≥ 16) hypoxemia.

We used the direct relationship between Spo_2 and Sao_2 to find the best Spo_2 aggregation. We compared mean and median function and different time windows from an exact time match to a 20-minute period before the Pao_2. We chose the aggregation method that minimized the bias.

Spo_2 Quality Assessment

In ICU, heart rate is continuously monitored using an electrocardiogram monitor and pulse oximeters. As both measurements should be identical in most clinical situations, the similarity is a way to validate the quality of the Spo_2 measures as already used by our team (35). We investigated if a higher difference in percentage between the pulse rate and the electrical heart rate (PR-HR) is associated with a greater MAE and lower ICC.

Empirical Optimization Model

As general equation by Hill (15) is complex to optimize by traditional mathematical techniques (**Fig. 1**, equation 1), we used an empirical method that calculated nearly all the combinations in the plausible ranges for Spo_2 (80–100%), P_{50} (20.0–40.0 mm Hg), m (0.950–0.999), and Hill (15) number n_H (2.50–3.50). Then, we identified the optimal combination that minimized the mean squared error (MSE) in the derivation dataset.

$$PaO_2 = \left(\frac{(P_{50})^{n_H}}{\frac{1}{SpO_2} - m} \right)^{\frac{1}{n_H}} \quad \text{(Equation 1)}$$

$$_{Hill}PaO_2 = \left(\frac{26^{2.7}}{\frac{1}{SaO_2} - 1} \right)^{\frac{1}{2.7}} \quad \text{(Equation 2 [reference 14])}$$

$$_{Severinghaus}SaO_2 = \left(\left(\frac{28.603^3}{(PaO_2)^3 + (150 \cdot PaO_2)} \right) + 1 \right)^{-1} \quad \text{(Equation 3 [reference 15])}$$

$$_{Gadrey}PaO_2 = \left(\frac{28.603^3}{\frac{1}{SpO_2} - 0.99} \right)^{\frac{1}{3}} \quad \text{(Equation 4 [reference 17])}$$

$$_{Sauthier}PaO_2 = \left(\frac{27.8^{2.8}}{\frac{1}{SpO_2} - 0.99} \right)^{\frac{1}{2.8}} \quad \text{(Equation 5)}$$

Figure 1. Known equations to estimate Pao_2 using Sao_2 or Spo_2.

This approach, also called exhaustive search, brute-force search, or grid search, is often used in machine learning hyperparameter optimization to solve a multidimensional problem (36). However, this optimization method requires a significant amount of time and computing resources.

Neural Network Models

In order to train the neural network model, we randomly split at the patient level the derivation cohort (BCH) into a train (85%) and a validation (15%) dataset. Once finalized, models were tested on the external distinct cohort (SJH). We used the last 20 minutes before Pao_2 to build a standardized multidimensional matrix of Spo_2, PR-HR (241 × 3, chronologically ordered). Data were normalized relatively to the minimum and maximum values of the training cohort. Missing values were imputed with the last value available. We compared three common types of networks able to assess serial or repeated measurements: multilayer perceptron, convolutional neural networks, and long short-term memory (LSTM) networks (37–39). Architecture, optimizer, and hyperparameters were adjusted to minimize the MSE.

Hypoxemia Severity Validation

Following the Pediatric Acute Lung Injury Consensus Conference (PALICC) (4), hypoxemia severity was defined using the OI: mild (values between 4 and 7.9), moderate (8–15.9), and severe (OI ≥ 16). If no arterial sample is available, PALICC guidelines recommend using the oxygen saturation index (OSI) where Pao_2 is replaced by Spo_2. OSI has different thresholds (mild 5–7.4, moderate 7.5–12.2, and severe ≥ 12.3) extrapolated from equations that estimate the OI (7). Mean airway pressures and Fio_2 were extracted at the exact same time the ABG was drawn and were the same for all models. We selected the best Pao_2 model to estimate the OI and compared its performance with the OSI on severity classification (kappa) and hypoxemia (mild and severe) discrimination (AUROC). Results were separated into two Spo_2 categories, below or equal to 97% and above 97%, even if OSI has not been validated in that upper range.

RESULTS

We included 64,926 Pao_2 (4,178 patients) with 46 million concomitant Spo_2 values. Patients and ABG characteristics in each cohort are described in **Table 1**. In Supplemental Digital Content (**SDC 1**, http://links.lww.com/CCX/A807), we showed that P_{50} varied from 20 to 40 mm Hg but was similar between age groups.

Aggregation Methods and Data Quality

We compared different time windows around the Sao_2 measurement (**SDC 2**, http://links.lww.com/CCX/A807) from –20 minutes to the exact time of the ABG. Although the differences were small, closer results to the exact time had a lower bias in both derivation and validation cohorts. Median function provided very slightly more biased results than the mean in both datasets.

Our results also showed that any difference between the pulse rate read by the pulse oximeter and the electrical heart rate was associated with a lower correlation and higher bias in the derivation cohort (**SDC 3**, http://links.lww.com/CCX/A807). The impact was more significant when the difference was greater than or equal to 3%. We confirmed the results in the validation cohort when the difference was less than 3% (ICC, 0.75; 95% CI, 0.74–0.75), greater than or equal to 3% (ICC, 0.44; 95% CI, 0.38–0.49), or unable to read a pulse (ICC, 0.49; 95% CI, 0.39–0.57). MAE increased from 2.2 (difference < 3%) to 3.4 (difference ≥ 3%) and 4.2 (no pulse read). ICC improved from 0.73 (95% CI, 0.72–0.73) to 0.75 (95% CI, 0.74–0.75; $p < 0.01$) after excluding those values, even if these results represented a minority of the measurements (51,777, 866, and 236, respectively, in the derivation cohort and 11,271, 582, and 194 in the validation cohort).

Empirical Optimization Model

Over the 1.015 million combinations of parameters using the equation structure by Hill (15), we found that results that minimized the MSE in the derivation cohort were P_{50} = 27.8 mm Hg, m = 0.99, and n_H = 2.8 (Fig. 1, equation 5). Performance in the derivation cohort is presented in **SDC 4** (http://links.lww.com/CCX/A807). Results for the validation cohort compared with the other models are shown in **SDC 5** (http://links.lww.com/CCX/A807) and illustrated in **Figures 2** and **3**. The $_{\text{Sauthier}}$estimated Pao_2 (ePao2) equation outperformed all existing models for ICC (0.38 vs 0.35 and lower, $p = 0.004$) and slopes (1.01 vs 1.14 and further from one; $p < 0.001$).

TABLE 1.
Patients and Arterial Blood Gases Characteristics

Variables	Boston Children's Hospital Derivation Cohort *n* (%) or median (IQR)	Sainte-Justine Hospital Validation Cohort *n* (%) or median (IQR)
Arterial blood gases, *n*	52,879	12,047
Patients (total), *n*	3,252	926
Adult patients (≥18), *n*	308 (9.5)	15 (1.6)
Arterial blood gases per patient	5 (2–17)	6 (2–13)
Age (yr)	2.4 (0.3–11.2)	1.4 (0.1–7.7)
Gender, female	23,545 (44.5%)	5,420 (45%)
pH	7.4 (7.3–7.4)	7.4 (7.4–7.5)
Pao_2 (mm Hg)	97 (75.4–128)	94.6 (72.6–129)
Spo_2 (%)	98 (96–100)	98.1 (95.5–100)
Lactate (mmol/L)	1.2 (0.8–1.8)	1.2 (0.8–2)
Gases under invasive ventilation	14,378 (27.2%)	7,547 (62.6%)
Mean airway pressure (cm H_2O)	11.4 (9–14.4)	10.5 (8.5–13)
Inspired oxygen fraction	0.4 (0.3–0.6)	0.4 (0.3–0.6)
Oxygenation index	4.5 (2.8–8.1)	4.3 (2.6–7.9)

IQR = interquartile range.

Neural Networks

The three networks had approximately similar performance (Fig. 2; and SDC 5, http://links.lww.com/CCX/A807), but the LSTM model was slightly less correlated with the observed Pao_2 than the others (ICC, 0.32 vs 0.35; $p < 0.01$). Bias was significantly lower ($p < 0.01$) for the neural networks compared with the mathematical equations (Fig. 2). The neural network structures are detailed in **SDC** (http://links.lww.com/CCX/A807).

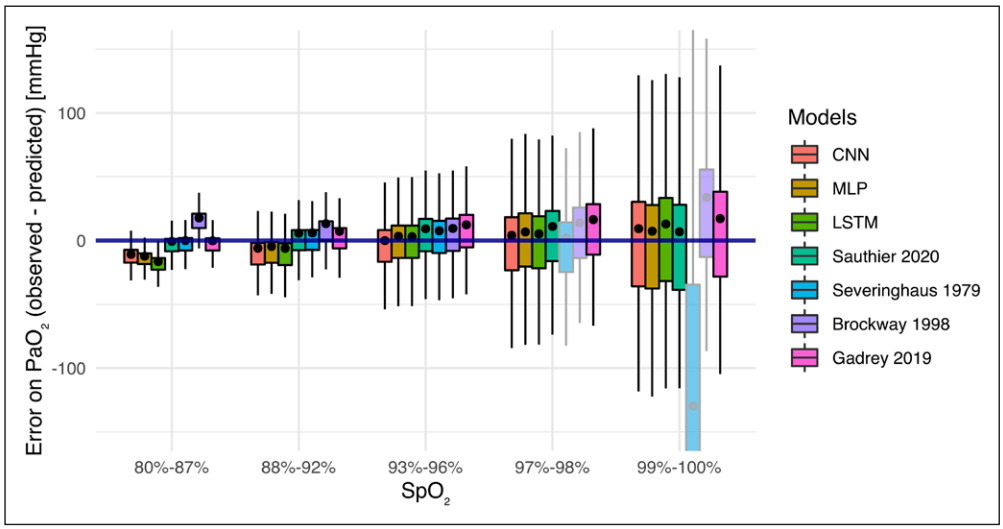

Figure 2. Grouped Bland-Altman plots showing models bias (mean of the errors, *black dots*) and limits of agreement (boxplot extremes) for different Spo_2 categories. Severinghaus (15) and Brockway and Hay (19) were not validated for $Spo_2 > 97\%$. CNN = convolutional neural network, LSTM = long short-term memory network, MLP = multilayer perceptron.

Hypoxemia Severity

Pediatric hypoxemia is determined by the OI or the OSI if no arterial sample is available. We compared the estimated OI based on $_{\text{Sauthier}}ePao_2$ with the OSI (**Table 2**). We used the same filtered Spo_2 by pulse rate and heart rate differences for both models. We found that eOI had better hypoxemia severity classification than the OSI for both Spo_2 less than or equal to 97% (Kappa 0.79 vs 0.60; $p < 0.001$, respectively) and

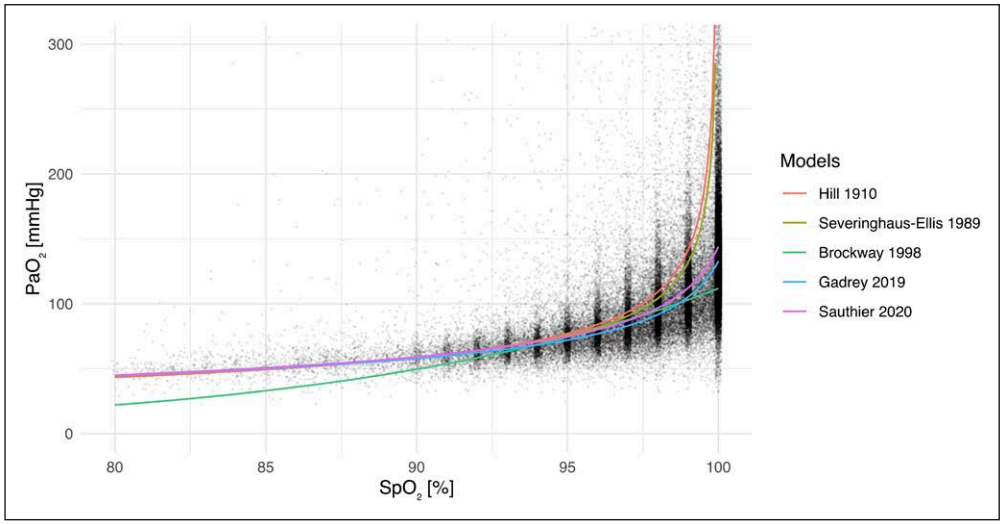

Figure 3. Pao_2 estimation on the derivation dataset using different models.

Spo_2 greater than 97% (Kappa 0.58 vs 0.52; $p < 0.001$, respectively). We also found that the eOI had a higher global agreement (ICC), lower fixed bias, higher correlation (slope), and better precision (limits of agreements difference and coefficient of determination) than OSI for both Spo_2 less than or equal to 97% and greater than 97%.

Furthermore, the $_{Sauthier}ePao_2$ had a similar classification performance (Kappa) for Spo_2 greater than 97% than OSI has for Spo_2 less than or equal to 97% (0.58 vs 0.60; $p = 0.06$). Discrimination for mild hypoxemia or more severe (OI ≥ 4) and for severe hypoxemia (OI ≥ 16) was similar for $ePao_2$ and OSI. Bland-Altman plots and AUROC plots are shown in **Figure 4** and **SDC 6** (http://links.lww.com/CCX/A807), respectively.

DISCUSSION

We have developed and validated a predictive model that can automatically and continuously estimate the Pao_2 using only noninvasive data that are commonly measured in ICU (Spo_2 and PR-HR) across two independent

TABLE 2.
Hypoxemia Severity Assessment

	$Spo_2 \leq 97\%$			$Spo_2 > 97\%$		
Metric (95% CI)	**OI Using OSI**	**OI Using $ePao_2$**	***p***	**OI Using OSI**	**OI Using $ePao_2$**	***p***
Weighted kappa	0.60 (0.58–0.62)	0.79 (0.77–0.81)	< 0.001	0.52 (0.50–0.54)	0.58 (0.56–0.59)	< 0.001
Intraclass correlation coefficient	0.91 (0.89–0.92)	0.93 (0.92–0.94)	0.029	0.62 (0.60–0.64)	0.72 (0.69–0.75)	< 0.001
Mean absolute error	3.68 (3.57–3.79)	1.90 (1.80–2.01)	< 0.001	3.58 (3.47–3.68)	1.60 (1.54–1.67)	< 0.001
Coefficient of determination (R^2)	0.84 (0.82–0.86)	0.87 (0.85–0.89)	0.043	0.51 (0.48–0.55)	0.55 (0.51–0.59)	0.170
Slope	0.81 (0.78–0.84)	1.02 (0.98–1.06)	< 0.001	0.42 (0.39–0.44)	0.93 (0.88–0.99)	< 0.001
Fixed bias	2.01 (1.85–2.17)	0.15 (0.02–0.28)	< 0.001	−0.48 (−0.63 to −0.34)	0.13 (0.05–0.21)	< 0.001
Proportional bias	−0.12 (−0.16 to −0.09)	0.09 (0.06–0.13)	< 0.001	−0.61 (−0.67 to −0.55)	0.26 (0.20–0.34)	< 0.001
Upper and lower limits of agreements difference	0.32 (0.30–0.34)	0.25 (0.23–0.28)	< 0.001	0.29 (0.28–0.31)	0.16 (0.15–0.18)	< 0.001
AUROC for mild hypoxemia (OI ≥ 4)	0.92 (0.91–0.93)	0.93 (0.92–0.94)	0.122	0.87 (0.86–0.88)	0.88 (0.87–0.89)	0.052
AUROC for severe hypoxemia (OI ≥ 16)	0.98 (0.97–0.98)	0.98 (0.98–0.99)	0.055	0.96 (0.95–0.97)	0.96 (0.95–0.97)	0.986

AUROC = area under the receiving operating characteristic curve, $ePao_2$ = estimated Pao_2, OI = oxygenation index, OSI = oxygenation saturation index.

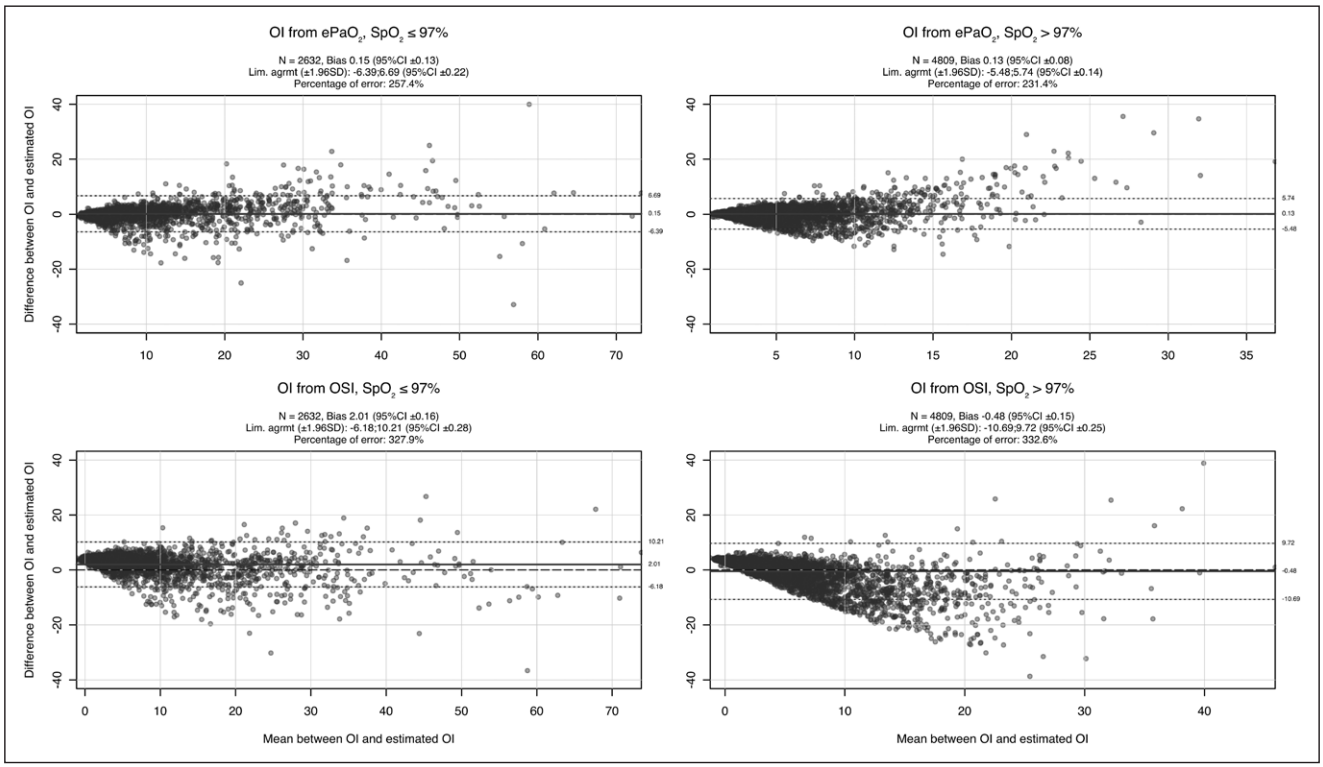

Figure 4. Bland-Altman plots for estimated oxygenation index (OI) using $ePao_2$ and oxygenation saturation index (OSI). $ePao_2$ = estimated Pao_2.

cohorts. A difference between the pulse read using the oximeter and the electrical heart rate was a good indicator of erroneous values, showing lower Sao_2-Spo_2 agreement. Spo_2 is optimally aggregated with a mean function on a 1-minute time window. A new mathematical equation ($_{Sauthier}ePao_2$) outperformed three existing mathematical equations for estimating Pao_2, and application of deep learning algorithms was unable to further improve performance. At Spo_2 values 99% and above, estimation of Pao_2 was still unbiased, but imprecise across all models. When $ePao_2$ is used to estimate OI, a better hypoxemia estimation is found than that with OSI. Although the improvement was modest, those findings are clinically relevant as they will allow building real-time clinical decision support systems that aim to automatically monitor degree of hypoxemia in patients on life support. Furthermore, OI and eOI will have the same thresholds, facilitating the interpretation. They will also allow continuous quantitative analysis of hypoxemia, such as duration and severity of each hypoxemia episode. Those results may also decrease blood draws and resource needs in a vulnerable population to anemia.

This study confirmed the contribution by Gadrey et al (17) to the equation by Severinghaus (15) and Hill (15) (SDC 5, http://links.lww.com/CCX/A807) that changing $m = 1$ to $m = 0.99$ improved most of the bias seen for Spo_2 values above 95%. However, we found no studies in the literature that tried to optimize Severinghaus (15) P_{50} value and n_H for a pediatric critical care population. In our study, these new parameters improved the agreement (ICC), the bias, and the slope, for all the clinically significant ranges of Spo_2. As we observed no clear association between age and P_{50}, it was unlikely that age would improve the model. However, precision remained low as a single value of Spo_2 can correspond to a broad range of Pao_2 (6). Future models may integrate patients' previous ABGs to improve continuous Pao_2 estimation precision. To improve reading in higher oxygenation values, other solutions such as the oxygen reserve index (40) used Svo_2 changes when Spo_2 is saturated. However, it requires specific proprietary material. To our knowledge, this study is the first to show different ways to retrospectively improve precision using continuous data: selecting a short time window aggregated by a mean function and exclude erroneous values when PR-HR is different.

In this study, we showed that three types of deep learning models were able to capture information from the continuous noninvasive input data resulting in the estimations with the lowest bias. However, the

difference with the $_{\text{Sauthier}}ePao_2$ equation was small (between 2 and 4 mm Hg) and is probably not clinically significant. Nonetheless, considering that a simple and an explicit mathematical equation can perform similarly, or better, neural networks do not seem to be worth the complexity they imposed, especially when ICC is low.

Our study has strengths. To our knowledge, this is the first and largest study to explore the Pao_2-Spo_2 relationship in the critical care field using only automatic continuously collected data. The ability to continuously assess hypoxemia severity is essential to improve early identification. Second, our model (Fig. 1, equation 5) is simple and explicit, and has been validated in a distinct external cohort. Even though most of the cohorts are children, the results did include adult patients (8%), and the $_{\text{Sauthier}}ePao_2$ parameters were close to the ones validated on adults, suggesting a good generalizability. Third, this study suggests that Spo_2 greater than 97% can still provide significant information regarding the hypoxemia severity, as also shown in severity scoring assessment (13). The $_{\text{Sauthier}}ePao_2$ had a similar classification performance (Kappa) for Spo_2 greater than 97% than OSI has for Spo_2 less than or equal to 97%. As precision is still lower for high Spo_2, it still supports the recommendation to minimize Fio_2 for appropriate Spo_2 targets (4).

However, our study has limitations. First, the retrospective nature of the study precludes the validation of the metadata (especially patient identification, timestamp, arterial labeling, and body location of the probe). However, those ABGs represented real-life situations, and both hospitals are committed to the higher standard of care including the validation of research databases (41). Second, measuring arterial oxygenation is a clinical decision and is not necessarily indicated for all patients. Thus, the patients included in this study are expected to be more severe than those who did not get any arterial measurement. However, overcoming this limitation in prospective trial would likely require invasive procedures in patients that are not warranted clinically. The large number of patients and ABG included in this study should mitigate this limitation. Third, we used a broad PICU population to derive and to validate the models. One would prefer to revalidate the model in a specific ARDS cohort before using it in a clinical decision support system dedicated to this entity.

CONCLUSIONS

In this two-center study with over 4,000 patients and nearly 65,000 ABGs paired with continuous data, we developed and validated a simple mathematical equation and methods to filter Spo_2 streams accurately and automatically in order to estimate the Pao_2 for patients admitted in critical care units using only continuous and noninvasive data. Despite a limited precision for high Spo_2, those results allow a precise OI estimation even for Spo_2 greater than 97% and are important for all future bedside clinical decision tools aiming at improving oxygenation titration in real time in critical care.

ACKNOWLEDGMENTS

We thank Dr. Shira Fischer, MD, PhD, for her advice and Dr. Craig Smallwood, RRT, PhD (deceased), for his assistance with the T3 dataset.

1 *Department of Anesthesiology, Critical Care and Pain Medicine, Boston Children's Hospital, Boston, MA.*

2 *Departments of Pediatrics, Sainte-Justine Hospital, Montreal, QC, Canada.*

3 *Departments of Anaesthesia and Pediatrics, Harvard Medical School, Boston, MA.*

4 *Innovation & Digital Health Accelerator, Boston Children's Hospital, Boston, MA.*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (http://journals.lww.com/ccejournal).

Dr. Sauthier acquired and analyzed the data, built the models, and drafted the article. All of the authors were responsible for designing the study, the interpretation of data, reviewing the article, approval of the version to be published, and providing agreement to be accountable for all aspects of the work.

Dr. Sauthier received funding from "Fonds de recherche du Québec en Santé" and the Université de Montréal (Faculty of Medicine), National Institutes of Health (R01AI084011 to Dr. Randolph), and the Centers for Disease Control and Prevention (to Dr. Randolph). Dr. Jouvet received funding from "Institut de Valorisation des données" for the development and analysis of the database at Sainte-Justine Hospital. The remaining authors have disclosed that they do not have any potential conflicts of interest.

Supported, in part, by public funding in the United States (to Dr. Randolph) and in Canada (to Drs. Sauthier and Jouvet).

The sponsor had no role in the design of the study, the collection and analysis of the data, or the preparation of the article.

For information regarding this article, E-mail: michael.sauthier@umontreal.ca

REFERENCES

1. Matthay MA, Zemans RL, Zimmerman GA, et al: Acute respiratory distress syndrome. *Nat Rev Dis Primers* 2019; 5:18
2. Chomton M, Brossier D, Sauthier M, et al: Ventilator-associated pneumonia and events in pediatric intensive care: A single center study. *Pediatr Crit Care Med* 2018; 19:1106–1113
3. Sauthier MS, Jouvet PA, Newhams MM, et al: Machine learning predicts prolonged acute hypoxemic respiratory failure in pediatric severe influenza. *Crit Care Explor* 2020; 2:e0175
4. Khemani RG, Smith LS, Zimmerman JJ, et al; Pediatric Acute Lung Injury Consensus Conference Group: Pediatric acute respiratory distress syndrome: Definition, incidence, and epidemiology: Proceedings from the Pediatric Acute Lung Injury Consensus Conference. *Pediatr Crit Care Med* 2015; 16:S23–S40
5. Bellani G, Laffey JG, Pham T, et al; LUNG SAFE Investigators; ESICM Trials Group: Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. *JAMA* 2016; 315:788–800
6. Eytan D, Mazwi ML, Goodwin AJ, et al: Revisiting oxygen dissociation curves and bedside measured arterial saturation in critically ill children. *Intensive Care Med* 2019; 45:1832–1834
7. Khemani RG, Thomas NJ, Venkatachalam V, et al; Pediatric Acute Lung Injury and Sepsis Network Investigators (PALISI): Comparison of SpO_2 to PaO_2 based markers of lung disease severity for children with acute lung injury. *Crit Care Med* 2012; 40:1309–1316
8. Riviello ED, Kiviri W, Twagirumugabe T, et al: Hospital incidence and outcomes of the acute respiratory distress syndrome using the Kigali modification of the Berlin definition. *Am J Respir Crit Care Med* 2016; 193:52–59
9. Tannoury JE, Sauthier M, Jouvet P, et al: Arterial partial pressures of carbon dioxide estimation using non-invasive parameters in mechanically ventilated children. *IEEE Trans Biomed Eng* 2021; 68:161–169
10. Pollard TJ, Johnson AEW, Raffa JD, et al: The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci Data* 2018; 5:180178
11. Mackle D, Beasley R, Bellomo R, et al: Conservative oxygen therapy during mechanical ventilation in the ICU. *N Engl J Med* 2020; 382:989–998
12. Helmerhorst HJ, Arts DL, Schultz MJ, et al: Metrics of arterial hyperoxia and associated outcomes in critical care. *Crit Care Med* 2017; 45:187–195
13. Slater A, Straney L, Alexander J, et al; Australian and New Zealand Intensive Care Society Paediatric Study Group and Centre for Outcomes and Resource Evaluation: The effect of imputation of PaO_2/FIO_2 from SpO_2/FIO_2 on the performance of the pediatric index of mortality 3. *Pediatr Crit Care Med* 2020; 21:520–525
14. Hill AV: The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves. *Proc Physiol Soc* 1910; 40:iv–vii
15. Severinghaus JW: Simple, accurate equations for human blood O_2 dissociation computations. *J Appl Physiol Respir Environ Exerc Physiol* 1979; 46:599–602
16. Ellis RK: Determination of PO_2 from saturation. *J Appl Physiol (1985)* 1989; 67:902
17. Gadrey SM, Lau CE, Clay R, et al: Imputation of partial pressures of arterial oxygen using oximetry and its impact on sepsis diagnosis. *Physiol Meas* 2019; 40:115008
18. Myburgh JA, Webb RK, Worthley LI: The P50 is reduced in critically ill patients. *Intensive Care Med* 1991; 17:355–358
19. Brockway J, Hay WW: Prediction of arterial partial pressure of oxygen with pulse oxygen saturation measurements. *J Pediatr* 1998; 133:63–66
20. Breuer HW, Groeben H, Breuer J, et al: Oxygen saturation calculation procedures: A critical analysis of six equations for the determination of oxygen saturation. *Intensive Care Med* 1989; 15:385–389
21. Pandharipande PP, Shintani AK, Hagerman HE, et al: Derivation and validation of Spo_2/Fio_2 ratio to impute for Pao_2/Fio_2 ratio in the respiratory component of the Sequential Organ Failure Assessment score. *Crit Care Med* 2009; 37:1317–1321
22. Rice TW, Wheeler AP, Bernard GR, et al; National Institutes of Health, National Heart, Lung, and Blood Institute ARDS Network: Comparison of the SpO_2/FIO_2 ratio and the PaO_2/FIO_2 ratio in patients with acute lung injury or ARDS. *Chest* 2007; 132:410–417
23. Khemani RG, Patel NR, Bart RD III, et al: Comparison of the pulse oximetric saturation/fraction of inspired oxygen ratio and the PaO_2/fraction of inspired oxygen ratio in children. *Chest* 2009; 135:662–668
24. Schwede M, Lee RY, Zhuo H, et al: Clinician recognition of the acute respiratory distress syndrome: Risk factors for under-recognition and trends over time. *Crit Care Med* 2020; 48:830–837
25. Sanz F, Dean N, Dickerson J, et al: Accuracy of PaO_2/FiO_2 calculated from SpO_2 for severity assessment in ED patients with pneumonia. *Respirology* 2015; 20:813–818
26. Johnson AE, Ghassemi MM, Nemati S, et al: Machine learning and decision support in critical care. *Proc IEEE Inst Electr Electron Eng* 2016; 104:444–466
27. Khemani RG, Rubin S, Belani S, et al: Pulse oximetry vs. PaO_2 metrics in mechanically ventilated children: Berlin definition of ARDS and mortality risk. *Intensive Care Med* 2015; 41:94–102
28. Leisman DE, Harhay MO, Lederer DJ, et al: Development and reporting of prediction models: Guidance for authors from editors of respiratory, sleep, and critical care journals. *Crit Care Med* 2020; 48:623–633
29. Collins GS, Reitsma JB, Altman DG, et al: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD statement. *Ann Intern Med* 2015; 162:55–63
30. Brossier D, Sauthier M, Mathieu A, et al: Qualitative subjective assessment of a high-resolution database in a paediatric intensive care unit-elaborating the perpetual patient's ID card. *J Eval Clin Pract* 2020; 26:86–91
31. Carpenter J, Bithell J: Bootstrap confidence intervals: When, which, what? A practical guide for medical statisticians. *Stat Med* 2000; 19:1141–1164
32. Liljequist D, Elfving B, Roaldsen KS: Intraclass correlation – a discussion and demonstration of basic features. *PLoS One* 2019; 14:e0219854
33. Shrout PE, Fleiss JL: Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull* 1979; 86:420–428

34. Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1:307–310
35. Ghazal S, Sauthier M, Brossier D, et al: Using machine learning models to predict oxygen saturation following ventilator support adjustment in critically ill children: A single center pilot study. *PLoS One* 2019; 14:e0198921
36. Chicco D: Ten quick tips for machine learning in computational biology. *BioData Min* 2017; 10:35
37. Ismail Fawaz H, Forestier G, Weber J, et al: Deep learning for time series classification: A review. *Data Min Knowl Discov* 2019; 33:917–963
38. Choi E, Schuetz A, Stewart WF, et al: Using recurrent neural network models for early detection of heart failure onset. *J Am Med Inform Assoc* 2017; 24:361–370
39. Catling FJR, Wolff AH: Temporal convolutional networks allow early prediction of events in critical care. *J Am Med Inform Assoc* 2020; 27:355–365
40. Szmuk P, Steiner JW, Olomu PN, et al: Oxygen reserve index: A novel noninvasive measure of oxygen reserve–a pilot study. *Anesthesiology* 2016; 124:779–784
41. Mathieu A, Sauthier M, Jouvet P, et al: Validation process of a high-resolution database in a paediatric intensive care unit-describing the perpetual patient's validation. *J Eval Clin Pract* 2021; 27:316–324