1. Introduction

The ratio of partial pressure of arterial oxygen (\(P_aO_2\)) to fraction of inspired oxygen (FiO\(_2\)), the ‘PF ratio’, quantifies severity of hypoxemic respiratory failure. A low value—low \(P_aO_2\) despite high FiO\(_2\)—indicates impaired lung function. PF ratios are, therefore, used in ICUs to measure illness severity within scores like the sequential organ failure assessment (SOFA) (Vincent et al 1996) and acute respiratory distress syndrome severity (Ranieri and ARDS Definition Task Force 2012) scores. The PF ratio had no well-defined role outside the ICU.
until 2016 when the operational criteria for sepsis changed from the systemic inflammatory response syndrome (SIRS) criteria (Sepsis-1 (Bone et al 1992) and Sepsis-2 (Levy et al 2001)) to organ failure criteria (Sepsis-3) (Singer et al 2016, Seymour et al 2016) and a rise in SOFA score of 2 in the setting of suspected infection became the operational definition for sepsis. The SOFA score rates respiratory failure based solely on PF ratios and therefore PF ratios now play a central role in the diagnosis of sepsis in acute-care patients. In non-intubated patients, the SOFA score is 0 for PF ratios over 400; 1 for PF ratios between 300 and 400, and 2 for PF ratios under 300. Accordingly, an infected patient with a PF ratio under 300 meets the definition of sepsis.

Calculating PF ratios requires both the knowledge of \( \text{FiO}_2 \) and \( \text{PaO}_2 \) measurement from an arterial blood gas (ABG). ABGs are infrequently obtained and therefore PF ratios are often unavailable in acute-care patients. When it is available, it is often a single measurement rather than periodic or continuous estimation which reduces clinical utility. This unavailability of PF ratios also hampers sepsis research. For example, in the Sepsis-3 study, PF ratios were missing in about 75% of non-ICU patients (Seymour et al 2016). The investigators addressed this problem with multiple imputation (Seymour et al 2016), which can be unreliable when large proportions of data are missing (Jakobsen et al 2017). Therefore, a strategy to optimally impute \( \text{PaO}_2 \) from the widely available oxygen saturations (\( \text{SpO}_2 \)) is needed.

Existing imputation techniques (figure 1, panel (A)) can be classified into those that model the oxygen-hemoglobin dissociation curve (Hill (Hill 1910, Goutelle et al 2008) and Severinghaus–Ellis (Severinghaus 1979, Ellis 1989)) and those that describe other (linear and log-log) relationships between \( \text{PaO}_2/\text{FiO}_2 \) and \( \text{SpO}_2/\text{FiO}_2 \) ratios (Rice et al 2007, Pandharipande et al 2009). The long and honourable history of relating the \( \text{SpO}_2 \) to the \( \text{PaO}_2 \) started in 1910, when the Nobel-prize-winning British biochemist A.V. Hill published a formulation relating the two as an example of the kinetics governing how enzymes bind and release their ligands. The general form of Hill’s equation is \( \text{SpO}_2 = \text{SpO}_2_{\text{max}} \times \left( \frac{\text{PaO}_2}{\frac{\text{PaO}_2}{\text{FiO}_2} + 1} \right)^{-n} \) where ‘\( \text{SpO}_2_{\text{max}} \)’ is the maximum possible value of \( \text{SpO}_2 \), ‘\( \text{PaO}_2 \)’ is that partial pressure of oxygen where hemoglobin is 50% saturated with oxygen and ‘\( n \)’ is the Hill’s coefficient which is an empirically determined constant for each enzyme-ligand system. The maximum possible value of \( \text{SpO}_2 \) is 1 and the adult human hemoglobin was thought by Hill to have a \( \text{PaO}_2 \) of 26 mmHg and a Hill coefficient of 2.7. Of note, these values were derived from laboratory solutions of hemoglobin rather than whole blood. By substituting these values, we arrive at the Hill equation and its inverse solution for \( \text{PaO}_2 \) which are depicted as equation (1) (figure 1, panel (A)). In 1979, Severinghaus made several modifications and showed that they improved performance at very low \( \text{SpO}_2 \) (<30%) in data from four healthy adults. He used a \( \text{PaO}_2 \) of 28.6025, a Hill coefficient of 3 (the exact number in his publication was 23 400 which is \( \approx 28.6025^3 \)) and added \( (150 + \text{PaO}_2) \) to the \( \text{PaO}_2^2 \) term. The addition of the \( (150 + \text{PaO}_2) \) term made the equation much harder than Hill’s to solve for \( \text{PaO}_2 \). However, in 1989, Ellis solved it. The resulting equations are depicted as equation (2) (figure 1, panel (A)). Then, in a departure from the prior approach of modelling the oxygen hemoglobin dissociation curve, a Vanderbilt University group took a model-independent approach. They fit straight lines, first to the plot of \( \text{PaO}_2/\text{FiO}_2 \) as a function of \( \text{SpO}_2/\text{FiO}_2 \) (equation (4): Rice et al 2007) and then to the plot of their logarithms (equation (5): Pandharipande et al).

Studies that compared these techniques (Sanz et al 2015, Brown et al 2016, Brown et al 2017) illustrated similarities between measured and imputed values using Spearman’s correlation coefficients (0.75 to 0.9, highest for Severinghaus–Ellis), ordinal score agreement (70% to 95%, highest for Severinghaus–Ellis) and areas under receiver operating characteristic curves for various PF ratio thresholds (0.77 to 0.96, highest for Severinghaus–Ellis and for lower thresholds). However, they did not compare biases or systematic disagreement patterns which are important measures of imputation quality (Kramer and Feinstein 1981, Ludbrook 1997, 2002). Our aim was to identify the optimal method for imputing \( \text{PaO}_2 \) from \( \text{SpO}_2 \), so that sepsis criteria may be accurately estimated without ABGs.

2. Materials and methods

2.1. Patient population and data collection

We retrospectively studied non-intubated in-patients from medical acute-care floors (general medicine, cardiology, gastroenterology, and oncology) of the University of Virginia Medical Center between 2013 and 2017 who had an \( \text{SpO}_2 \) value recorded no more than 10 min prior to an ABG draw and who had information sufficient to estimate the \( \text{FiO}_2 \) (oxygenation device, flow rate and set \( \text{FiO}_2 \)) recorded within 12 hours prior to the ABG draw. Because model-based imputation equations (Hill, Severinghaus–Ellis) do not impute PF ratios at \( \text{SpO}_2 \) of 100%, we only compared cases with saturations < 100% (n = 492). We calculated measured PF ratios using the \( \text{PaO}_2 \) from an ABG and the \( \text{FiO}_2 \) (\( \text{FiO}_2 \) was estimated using conventional methods shown in figure 2). We obtained four sets of imputed PF ratios by applying the existing imputation equations to the \( \text{SpO}_2 \) level and \( \text{FiO}_2 \). From these measured and imputed PF ratios, we derived measured and imputed SOFA respiratory component scores.
The final data set contained 492 instances with one set of ‘true PF ratios’ and ‘true SOFA scores’ from ABGs, and four sets of ‘imputed PF ratios’ and ‘imputed SOFA scores’ from SpO₂.

We identified data using queries to the university’s enterprise data warehouse. The accuracy of these queries was iteratively optimized by performing quality checks through random manual chart reviews until they were...
error-free. We performed all data processing and analyses using R version 3.5.1 (R Core Team 2018). The study was approved with waiver of informed consent by University of Virginia’s Institutional Review Board for Health Sciences Research (Protocol 20677).

2.2. Analysis techniques

We assayed the goodness of fit of the equations by plotting the imputed PF ratio as a function of the observed PF ratio and measuring both the Spearman correlation coefficient \( r \) and the proportional and fixed biases of the Model 2 (major axis) regression line. Proportional bias is the deviation of the regression line slope from 1 and fixed bias is the deviation of the regression line intercept from 0 (Ludbrook 1997). We assayed SOFA respiratory component score imputation by measuring the frequency at which the imputed SOFA value differed from the true value and whether the difference resulted in an underestimate or overestimate of hypoxemia severity. We tested for ordinal score imputation bias using a modified McNemar test which evaluated the null hypothesis that there was no bias in ordinal score imputation (Kramer and Feinstein 1981, Ludbrook 2002). This test calculates \( \chi^2 \) using the formula \( \left( N_{UE} - N_{OE} \right) / \left( N_{UE} + N_{OE} \right) \) where \( N_{UE} \) is the number of underestimates and \( N_{OE} \) is the number of overestimates. \( \chi^2 \) values larger than 3.841 (corresponding with \( p \)-values of under 0.05 with 1 degree of freedom) indicate statistically significant bias.

2.3. Deriving a modified model based equation

We hypothesized that biases in imputation from model-based equations were attributable to the \( \left( \frac{1}{\text{SpO}_2} - 1 \right) \) term in their denominators. As \( \text{SpO}_2 \) approaches 1, the denominator approaches 0 and causes a rise in imputed \( \text{P}_2\text{O}_2 \) that exceeds the rise seen in real life. Figure 4, panel (A) clearly shows this effect. To remedy this singularity, we tested modifications of the \( \left( \frac{1}{\text{SpO}_2} - 1 \right) \) term in the model-based equations in the form of \( \left( \frac{1}{\text{SpO}_2} - n \right) \). Plotting these variants made it clear that the model that best described the real-world observations would be one where \( 0.95 < n < 1 \) (i.e. to the left of blue line but to the right of the pink line in figure 4, panel (A)). Thus, we created 51 distinct model based equations where ‘\( n \)’ in \( \left( \frac{1}{\text{SpO}_2} - n \right) \) was varied between 0.950 and 0.999, and the resulting equations were assayed using same performance metrics described above. To avoid the possibility that a modifier showed superior performance just by chance (i.e. over-fit to our dataset), we performed a rigorous bootstrap resampling analysis for internal cross-validation. We measured the performance metrics for each modifier in 200 bootstrapped 70% random samples of our data and then averaged them before comparison. This modified equation with the best performance was then compared to existing equations.
3. Results

3.1. Characteristics of the study population
Of the 492 patients included in the study, the median age was 67 (interquartile range: 59–75); 58% were men, 88% were white and 12% were black.

3.2. Comparing imputed and measured PF ratios
Severinghaus–Ellis equation based imputation performed better overall than imputation using the Hill, Rice, and Pandharipande equations. However it underestimated the SOFA respiratory component score at a frequency of 17%, and there were significant degrees of proportional (slope 1.3), fixed (intercept −46.3) and ordinal imputation ($\chi^2$ of 36.1; $p$-value < 0.001) biases (figure 3 and table 1).

3.3. Deriving an optimal model-based equation
Much of the bias associated with the Hill and Severinghaus–Ellis equations resulted from inaccuracies at high SpO2. When observations with SpO2 > 96% were removed, the bias substantially declined but remained statistically significant (table 1). To test whether these findings arose from the singularity in the model-based equations when SpO2 = 1, we replaced the $[(1/\text{SpO2}) - 1]$ term with $[(1/\text{SpO2}) - n]$ where $n$ was varied between 0.950 and 0.999. Figure 4 shows that the modifiers between 0.987 and 0.993 were free of bias. Of the modifiers we tested, 0.99 was the optimum replacement based on improved proportional (slope 1.0), fixed (intercept −9.6) and ordinal imputation ($\chi^2$ of 0.13) biases with preserved correlation coefficient (0.82). The frequency of underestimation of the SOFA respiratory component score was reduced to 7% and accuracy rose to 86%. As described before, these results were obtained after rigorous bootstrap resampling analysis for internal cross-validation, thereby reducing probability of over-fit to our particular dataset.

As suggested by Severinghaus (1979), we confirmed that the impact of the $(150 + P_{\text{aO}_2})$ term in equation (2) (figure 1) was only observable at very low SpO2 (<30%) and not within the clinically significant range of SpO2 (80%–99%). Therefore, eliminating it made the inverse solution less complex without affecting imputation quality. Based on these findings, we concluded that the optimal equation for imputation of PF ratios and SOFA scores among non-intubated acute care floor patients was equation (5), $P_{\text{aO}_2} = \left(\frac{23,400}{\text{SpO}_2 - 0.99}\right)^{\frac{1}{3}}$, which outperformed all existing model-based and model-independent imputation equations (table 1; figures 1 and 3).

3.4. Sensitivity analyses
In our primary analysis, we only used ABGs that were accompanied by a documented SpO2 no more than 10 min prior to ABG draw time. To test whether this window impacts the results, we varied the allowable lag between ABG and SpO2 from 1 min ($n = 107$) to 30 min ($n = 878$) and the findings did not change. Analysis on a subset of African American subjects ($n = 57$) was performed given some past studies suggesting that higher SpO2 targets were needed to achieve same PaO2 levels in black patients (Jubran and Tobin 1990). This analysis showed similar results as entire data set.

3.5. Imputing $P_{\text{aO}_2}$ at 100% SpO2
Unlike the Hill or Severinghaus–Ellis model-based equations, our modified equation allowed for an imputed PF ratio at SpO2 values of 100%. In a set of 111 ABGs where the nearest SpO2 reading was 100%, the imputation of our modified equation remained unbiased in the continuous form, but had statistically significant bias towards underestimation with ordinal scores. The degree of linear model fit also declined considerably ($r = 0.58$). This led to a lower frequency of accurate SOFA score imputation (50.5%) and a rise in the frequency of both underestimation (35%) and overestimation (14%).

4. Discussion

4.1. Major findings and clinical applicability
We studied various methods of imputing PF ratios and the respiratory components of the SOFA scores from SpO2 and FiO2. We focused on the systematic patterns of clinically significant errors (bias, underestimation) that resulted from each method so that a method with optimal clinical utility could be identified. This included studying the performance of a modified model-based equation that we derived by building on past work. The primary goal of this study was to optimize the estimated SOFA respiratory component score by optimally imputing the PF ratio from the SpO2 and FiO2.

We found that all model-based equations (Hill, Severinghaus–Ellis, and our modified equation) showed better model fit ($r = 0.80–0.82$) than the model-independent equations (Rice and Pandharipande; $r = 0.75$). This
affirmed the fundamental importance of the biochemical model of oxygen dissociation from hemoglobin in estimating $P_aO_2$ from $SpO_2$.

While our results showed that the Severinghaus–Ellis equation outperformed all other existing equations, we found that it systematically underestimated hypoxemia severity (i.e. overestimated PF ratios). Because failure of early detection of respiratory failure is potentially catastrophic, we sought to derive an equation that would improve the accuracy of SOFA score imputation by minimizing underestimates of hypoxemia severity. We extended the work of Hill, Severinghaus, and Ellis to derive a modified model-based equation which had unbiased imputation at all $SpO_2$ levels. The general form of our equation is the same as the Hill equation rather than the Severinghaus–Ellis equation because we determined that the ‘$150 \times P_aO_2$’ term added by Severinghaus–Ellis only affected imputation at very low $SpO2$ (<30%) and not in the clinically relevant range of 85% to 99%.

Figure 3. Comparing the imputation performance of the equations. Each panel of this figure shows the performance of one equation through a scatter plot (imputed PF ratio as a function of measured PF ratio) and a confusion matrix (frequency of agreements and disagreements between measured and imputed SOFA scores). The top two panels (A) and (B) deal with the existing model-based equations (Hill and Severinghaus–Ellis respectively) and the bottom two panels (D) and (E) deal with the model-independent equations (Rice and Pandharipande respectively). The larger central panel (C) shows the performance of the modified model-based equation that we are introducing in this paper. In the scatter plots, the red line is a model (major axis) regression line and the blue dotted lines are 95% confidence interval lines. The green line is the identity line (slope 1, intercept 0) which is what the regression line looks like in the event of ideal, bias-free imputation. In all panels except panel (C), the regression line deviates significantly from the identity line whereas in panel (C) the regression line and the identity line are almost indistinguishable. This demonstrates that all existing equations produce significantly biased imputations, a problem that is eliminated by our modified model-based equation. In the confusion matrices, the diagonal comprised of the green boxes shows the frequency of instances where there was agreement between the measured and imputed SOFA scores. The red boxes (on top and to the left of the green diagonal) represent the instances were imputed SOFA scores were an underestimation of the true SOFA scores and the yellow boxes (on the bottom and to the right of the green diagonal) represent the instances were imputed SOFA scores were an overestimation of the true SOFA scores. Panels (A), (B) and (D) show a heavy bias towards underestimates whereas panel (E) shows a heavy bias towards overestimation. Panel (C) shows that our modified model based equation shows no bias with essentially equal number of underestimates and overestimates while retaining high accuracy.
However, we retained the values of $P_{50}$ and the Hill coefficient ‘$n$’ (28.6025 and 3 respectively) as determined by Severinghaus because these gave better imputation performance in our data. Then, as shown in figure 4, we eliminated the singularity common to both the Hill and Severinghaus–Ellis equations by replacing \((1/\text{SpO}_2) - 0.99\) with \(1/\text{SpO}_2 - 1\). This final change, our most important contribution, was responsible for eliminating the bias and underestimation seen with the Hill and Severinghaus–Ellis equations. Although the bias in these equations was most prominent at high \text{SpO}_2 levels (>96%), its presence adversely impacted the overall utility of the equation. Only in the narrow use case of a bedside spot-check is it possible to work around this flaw. In this case, one can wean all supplemental oxygen before imputing a PaO$_2$ and in cases where from \text{SpO}_2 remains > 96% off supplemental oxygen, one can assume that there is no respiratory dysfunction present with regard to SOFA scoring. However, in most clinical and research use cases, users are conducting a retrospective review of PF ratios rather than a spot check at the bedside (a clinician trying to determine if hypoxemia has worsened in the last 12 h or a researcher determining PF ratios in a retrospective cohort of infected subjects). For such applications, weaning before imputation is obviously not feasible and omitting data points is not optimal (\text{SpO}_2 > 96% are often up to half of the available data on acute care floors; 45% in our data).

The only solution in such cases is to use an imputation equation and biased imputation at any level of \text{SpO}_2 limits the utility of imputation at every level of \text{SpO}_2. For example, if a patient is weaned from 4 liters per minute by nasal cannula (FiO$_2$ ~ 0.33) to room air and their \text{SpO}_2 drops from 98% to 90%, the Severinghaus–Ellis equation would suggest that the PF ratio has approximately dropped from 316 to 281 suggesting acutely deteriorating oxygenation (a rise in SOFA from 1 to 2). However, our modified equation correctly recognizes that the PF ratio

### Table 1. Imputation performance of model-based and model-independent equations.

|                  | Metrics of PF ratio imputation | Metrics of SOFA (pulmonary) score imputation |
|------------------|--------------------------------|---------------------------------------------|
|                  | Correlation coefficient ($r$) | Regression slope; (95% CI) | Regression intercept; (95% CI) | Accuracy N(%) | Underestimation N(%) | Overestimation N(%) | Bias $\chi^2$ (p-value) |
| **Model-based equations** |                                 |                              |                               |               |                   |                     |                          |
| Hill             | 0.80                           | 1.4; (1.3 to 1.5)            | -65; (86 to -46)              | 372           | 101               | 19$^a$              | 56                   |
| Severinghaus–Ellis |                                |                               |                               |               |                   |                     | (<0.001)             |
| All \text{SpO}_2 | 0.81                           | 1.3; (1.2 to 1.4)            | -46.3; (65 to -29)            | 389           | 82                | 21                  | 36.1                 |
| $\text{SpO}_2 \leq 0.96$ | 0.85                           | 1.1; (1.0 to 1.2)            | -15.8; (31 to -2)             | 253           | 11                | 5                   | 2.25                 |
| $\text{SpO}_2 = 97-99$ | 0.70                           | 1.4; (1.2 to 1.6)            | -51.5; (103 to -8)            | 136           | 71                | 16                  | 34.77                |
| $\text{SpO}_2 = 100$ | Function not defined           |                               |                               |               |                   |                     | (<0.001)             |
| **New equation** |                                 |                               |                               |               |                   |                     |                          |
| All \text{SpO}_2 | 0.82$^a$                       | 1.0$^b$; (0.9 to 1.1)        | -9.6$^b$; (24 to 4)           | 425$^a$       | 35                | 32                  | 0.13$^b$             |
| $\text{SpO}_2 \leq 0.96$ | 0.85                           | 1.0; (0.9 to 1.1)            | -10.1; (24 to 3)              | 253           | 8                 | 8                   | 0                    |
| $\text{SpO}_2 = 97-99$ | 0.71                           | 1.0; (0.9 to 1.1)            | -3.6; (33 to 33)              | 172           | 27                | 24                  | 0.17                 |
| $\text{SpO}_2 = 100$ | 0.58                           | 1.1; (0.9 to 1.5)            | -9.2; (122 to 74)             | 56            | 39                | 16                  | 9.6                  |
| **Model-independent equations** |                               |                               |                               |               |                   |                     |                          |
| Rice             | 0.75                           | 1.4; (1.3 to 1.5)            | -70.9; (95 to -49)            | 379           | 83                | 30                  | 24.86                |
| Pandharipande    | 0.75$^a$                       | 0.7; (0.6 to 0.8)            | 72.5; (61 to 84)              | 425$^a$       | 19$^a$            | 48                  | 12.55                |

$a$ Best performance among all equations for that metric.  
$b$ No statistically significant bias was detected.
is stable around 275 and that the true SOFA score is 2 in both scenarios. Thus, an unbiased imputation equation represents a crucial advance in improving the reliability of imputed PF ratio values for use in research and clinical practice. For ease of incorporation into practice, especially in resource poor settings, we have created a convenient look-up table (figure 5).

Finally, we found that while our modified model-based equation allowed for imputation at SpO2 of 100%, the model fit at this level was quite poor and led to high levels of inaccuracy when imputing SOFA scores. This problem is related to some unique biochemical implications of the common clinical finding of 100% saturation which occurred in 19% of patients in our study—all of whom were receiving supplemental oxygen with a median FiO2 0.35. Complete saturation of hemoglobin is achieved at PaO2 around 120 mmHg, a value far below the maximum achievable values. For example, the maximum achievable Pao2 values are 195 mmHg at FiO2 of 0.35, the median FiO2 in our series, and 659 mmHg at FiO2 of 1, the maximum FiO2 in our series (maximum achievable

---

**Figure 4.** Deriving an optimally modified model-based equation which provides bias-free imputation. Panel (A) shows the impact of modifying the \( \left( \frac{1}{\text{SpO2}} - 1 \right) \) term. The black dots are true PaO2 and SpO2 values measured from ABG and pulse oximetry respectively. The blue line shows the imputation pattern of a model-based equation where the \( \left( \frac{1}{\text{SpO2}} - 1 \right) \) term has been left unchanged. It clearly shows that the rise in imputed PaO2 with rising SpO2 far exceeds the rise in measured PaO2. The green and pink lines show the imputation patterns of model-based equations where that term was replaced with \( \left( \frac{1}{\text{SpO2}} - 0.99 \right) \) and \( \left( \frac{1}{\text{SpO2}} - 0.95 \right) \) respectively. While the larger plot shows imputation performance at all SpO2, the white rectangle demarcates the range of SpO2 most commonly encountered in clinical practice (85% to 100%) and a zoomed in plot of this area is shown in the bottom right corner. Panels (B)–(D) show the impact of these modifications on SOFA score imputation, proportional bias and fixed bias respectively. Each point on the x-axis of these panels is a unique ‘n’ inserted into the denominator term \( \left( \frac{1}{\text{SpO2}} - n \right) \). Thus, each x-axis point is a distinct model-based equation. The y-axis represents the performance metrics for the distinct equations being compared (percentage of agreement/disagreement in Panel (B), slope of major axis regression line and its 95% CIs in panel (C) and intercept of major axis regression line and its 95% CIs in panel (D)). To avoid the possibility of over fit to our particular dataset, we used 200 bootstrapped samples for cross-validation. Therefore, for each distinct equation on the x-axis there are 200 unique performance metric values (coloured dots) and 1 mean performance value (black line). In each of these panels the rectangle demarcates the modifier range which was free of any significant bias. In panel (B), the green, red and yellow dots represent the percentage of imputed SOFA scores that were accurate estimates, under-estimates and over-estimates respectively of the true SOFA. Compared to a model-based equation with \( \left( \frac{1}{\text{SpO2}} - 1 \right) \) in the denominator (solid black vertical line), the variant with \( \left( \frac{1}{\text{SpO2}} - 0.99 \right) \) in the denominator (dotted black vertical line) has higher accuracy (86.4% versus 79.1%; \( p < 0.05 \)), lower underestimates (7.1% versus 16.7%; \( p < 0.05 \)) and higher overestimates (6.5% versus 4.2%; \( p < 0.05 \)). In panel (C), the pink and blue dots represent the slope of the major axis regression line and its 95% confidence intervals respectively. Compared to a model-based equation with \( \left( \frac{1}{\text{SpO2}} - 1 \right) \) in the denominator (solid black vertical line), the variant with \( \left( \frac{1}{\text{SpO2}} - 0.99 \right) \) in the denominator (dotted black vertical line) has significantly less proportional bias (slope of 1.0 [95% CI: 0.9–1.1] versus 1.3 [95% CI: 1.2–1.4], \( p < 0.05 \)). In panel (D), the pink and blue dots represent the intercept of the major axis regression line and its 95% confidence intervals respectively. Compared to a model-based equation with \( \left( \frac{1}{\text{SpO2}} - 1 \right) \) in the denominator (solid black vertical line), the variant with \( \left( \frac{1}{\text{SpO2}} - 0.99 \right) \) in the denominator (dotted black vertical line) has significantly less fixed bias (intercept of −9.6 [95% CI: −24 to 4] versus −46 [95% CI: −65 to −29], \( p < 0.05 \)).
levels calculated by assuming normal values in the alveolar gas equation and a normal A-a gradient—see table 2. Thus, the \( P_{\text{a}O_2} \) at 100% \( O_2 \) saturation cannot be determined by imputation.

Sepsis remains a major problem on acute-care wards, and our method of calculating the respiratory component of the SOFA score may improve the ease and accuracy of diagnosis and clinical decision-making. It is important to note that the value we selected for our modified constant was chosen specifically to prioritize avoiding underestimates at the cost of an increase in overestimates. Other clinicians may favor a different tradeoff, and thus may select a different modifier within the bias free zone of 0.987 to 0.993 instead of the \( (1/\text{SpO}_2) - 0.99 \) that we favor. Additionally, our sample may exhibit a selection bias towards true hypoxemia. For these reasons, our equation is likely to be more sensitive than specific thereby identifying a higher proportion of all true hypoxemia cases at the expense of an increase in false positives. Therefore, the most appropriate use of our modified equation would be as a sensitive screening or risk stratification tool for the early detection of respiratory failure due to sepsis, pulmonary edema, pneumonia, or pulmonary embolism. An alert for hypoxemia generated using our equation should be supported by clinical assessment and could be confirmed with ABG testing. It must be noted, of course, that the utility in screening for respiratory failure only applies to hypoxemic respiratory failure syndromes and not for hypercapnic respiratory failures or other complex acid-base derangements which may impact respiratory status. For these cases, ABG will remain indispensable.

### Table 2. Estimating maximum achievable \( P_{\text{a}O_2} \) at a particular \( \text{FiO}_2 \) level for our subjects.

\[
\begin{align*}
P_{\text{a}O_2} &= (\text{FiO}_2 \times (\text{P}_{\text{atm}} - \text{P}_{\text{H}_2\text{O}})) - (\text{P}_{\text{a}CO_2} R) \\
\text{P}_{\text{a}O_2} &\text{max} = (\text{P}_{\text{a}O_2}) - (A-a \text{ Gradient})_{\text{min}} \\
\text{P}_{\text{a}O_2}\text{max}(\text{min}) &\text{maximum possible arterial partial pressure of oxygen at a certain \( \text{FiO}_2 \);} \\
\text{P}_{\text{a}O_2} &\text{alveolar partial pressure of oxygen at that \( \text{FiO}_2 \);} \\
A-a \text{ Gradient}(\text{min}) &\text{minimum possible gradient between \( P_{\text{a}O_2} \) and \( P_{\text{a}O_2} \)} \\
\text{By definition, the A-a gradient is the difference between} &\text{Thus, the maximal possible \( P_{\text{a}O_2} \) occurs when A-a gradient is} \\
\text{\( P_{\text{a}O_2} \) and \( P_{\text{a}O_2} \).} &\text{at its minimum possible value which is 5 mmHg.} \\
\text{Assuming the minimum possible A-a gradient and with normal values} &\text{Assuming the minimum possible A-a gradient and with normal values in the alveolar gas equation as above, the maximum achievable} \\
\text{in the alveolar gas equation and a normal A-a gradient—see table 2.} &\text{value of \( P_{\text{a}O_2} \) at a given \( \text{FiO}_2 \) is:} \\
\text{Thus, the \( P_{\text{a}O_2} \) at 100% \( O_2 \) saturation cannot be determined by imputation.} &\text{P}_{\text{a}O_2}\text{max}(\text{min}) = (\text{FiO}_2 \times 714) - 55
\end{align*}
\]

### Figure 5. This figure serves as a convenient way to determine PF ratios (and SOFA scores) associated with common levels of \( \text{FiO}_2 \) and \( \text{SpO}_2 \). (lpm = liters per minute; NC = nasal cannula; FM = face mask; NRB = non-rebreather).

![Figure 5](image_url)

| Flow/Device | \( \text{FiO}_2 \) | 88% | 89% | 90% | 91% | 92% | 93% | 94% | 95% | 96% | 97% | 98% | 99% |
|-------------|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Room Air    | 21%             | 258 (2) | 266 (2) | 275 (2) | 285 (2) | 296 (2) | 309 (1) | 325 (1) | 343 (1) | 366 (1) | 395 (1) | 436 (0) | 501 (0) |
| 1 lpm by NC | 24%             | 226 (2) | 233 (2) | 241 (2) | 250 (2) | 259 (2) | 271 (2) | 284 (2) | 300 (1) | 320 (1) | 346 (1) | 382 (1) | 438 (0) |
| 2 lpm by NC | 27%             | 201 (2) | 207 (2) | 214 (2) | 222 (2) | 231 (2) | 241 (2) | 253 (2) | 267 (2) | 284 (2) | 307 (1) | 339 (1) | 390 (1) |
| 3 lpm by NC or 30% by FM | 30% | 183 (3) | 187 (3) | 193 (3) | 200 (2) | 208 (2) | 217 (2) | 227 (2) | 240 (2) | 256 (2) | 277 (2) | 305 (1) | 351 (1) |
| 4 lpm by NC | 33%             | 164 (4) | 170 (3) | 175 (3) | 182 (2) | 189 (2) | 197 (2) | 207 (2) | 218 (2) | 233 (2) | 252 (2) | 278 (2) | 319 (2) |
| 5 lpm by NC or 35% by FM | 35% | 155 (5) | 160 (3) | 165 (3) | 171 (3) | 178 (3) | 186 (3) | 195 (3) | 206 (2) | 219 (2) | 237 (2) | 262 (2) | 301 (2) |
| 6 lpm by NC or 40% by FM | 40% | 136 (6) | 140 (4) | 145 (3) | 150 (3) | 156 (3) | 162 (3) | 170 (3) | 180 (3) | 192 (3) | 207 (2) | 229 (2) | 263 (2) |
| 8 lpm by NC or 45% by FM | 45% | 121 (7) | 124 (3) | 128 (3) | 133 (3) | 138 (3) | 144 (3) | 152 (3) | 160 (3) | 171 (3) | 184 (3) | 204 (2) | 234 (2) |
| 10 lpm by NC or 50% by FM | 50% | 109 (8) | 112 (3) | 116 (3) | 120 (3) | 125 (3) | 130 (3) | 136 (3) | 144 (3) | 154 (3) | 166 (3) | 183 (3) | 210 (2) |
| 15 lpm by NC or 60% by FM | 60% | 90 (10) | 93 (4) | 96 (4) | 100 (3) | 104 (3) | 108 (3) | 114 (3) | 120 (3) | 128 (3) | 138 (3) | 153 (3) | 175 (3) |
| 100% by NRB mask | 100% | 54 (14) | 56 (4) | 58 (4) | 60 (4) | 62 (4) | 64 (4) | 68 (4) | 72 (4) | 74 (4) | 83 (4) | 92 (4) | 105 (3) |

Note: This table shows PF ratios (and SOFA scores) associated with common \( \text{FiO}_2 \) and \( \text{SpO}_2 \) values. For non-intubated patients, replace all SOFA scores of 3 and 4 with 2.
4.2. Strengths and weaknesses
Our study rigorously quantified systematic errors of bias and inaccurate ordinal imputations that occur when using existing or our derived equation. These errors are among the most important considerations when comparing methods of imputation (Kramer and Feinstein 1981, Ludbrook 1997, 2002). Additionally, our results were based on observations from 492 non-intubated acute care patients which provide confidence in the ability to estimate PF ratios for sepsis diagnosis in similar patients. Moreover, because the data used to formulate past equations were either from young healthy male subjects (Hill and Severinghaus) or data from mechanically ventilated ICU patients (Rice and Pandharipande), our work is likely to work well in the acute-care setting where imputation could be of tremendous value given sparing use of ABGs.

We studied cases where physicians deemed it clinically necessary to obtain an ABG outside the ICU. It is therefore plausible that the patients in our dataset were more likely to have true hypoxemia. As such, the results of our study are most generalizable to patients in whom an acute illness that predisposes patients to acute respiratory failure is suspected. Further research is needed to determine suitability of our imputation strategy for use in settings when an acute illness is not suspected.

The bootstrap cross-validation analysis that we used is a standard technique for validating results in studies of single datasets and our use of 200 bootstrapped samples is sufficiently rigorous (Collins et al 2015). Still, prospective external validation of our equation in disparate datasets would be very desirable.

Finally, increasing inaccuracy at high SpO2 was a limitation of all imputations. We attributed this to the wide range of PaO2 corresponding to each SpO2 level in this range. Since oxygen dissolves into blood and cannot be measured from hemoglobin saturation at high PaO2, the same degree of imprecision in SpO2 measurement techniques affected model fit more adversely at higher saturations.

5. Conclusion
We propose a modified oxygen-hemoglobin dissociation model-based equation for imputing PF ratios from SpO2 and FiO2 in non-intubated acute-care patients. Being the first bias-free equation, it would allow for wider clinical use of imputed PF ratios and could improve the performance of early warning and risk stratification scores used in acute-care patients.

Conflicts of interest
No conflicts exist for any of the specified authors.

Funding information
We have no sources of funding or support to disclose.

Notation of prior abstract publication/presentation
Gadrey S et al. IMPUTING THE RATIO OF PARTIAL PRESSURE OF ARTERIAL OXYGEN TO FRACTION OF INSPIRED OXYGEN: THE OPTIMAL STRATEGY IN NON-INTUBATED FLOOR PATIENTS. Abstract published at Hospital Medicine 2019, March 24–27, National Harbor, Md. Abstract 122. https://www.shmabstracts.com/abstract/imputing-the-ratio-of-partial-pressure-of-arterial-oxygen-to-fraction-of-inspired-oxygen-the-optimal-strategy-in-non-intubated-floor-patients/ 2019.

References
ARDS Definition Task Force and Ranieri V M et al 2012 Acute respiratory distress syndrome: the Berlin Definition JAMA 307 2526–33
Bone R C et al 1992 Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine Chest 101 1644–55
Brown S M et al 2016 Nonlinear imputation of Pao2/Fio2 From SpO2/FIO2 among patients with acute respiratory distress syndrome Chest 150 307–13
Brown S M et al 2017 Nonlinear imputation of PaO2/FIO2 From SpO2/FIO2 among mechanically ventilated patients in the ICU: a prospective, observational study Crit. Care Med. 45 1317–24
Collins G S, Reitsma J B, Altman D G and Moons K G M 2015 Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement Ann. Intern. Med. 162 55
Ellis R K 1989 Determination of PO2 from saturation J. Appl. Physiol. 67 902
Goutelle S et al 2008 The Hill equation: a review of its capabilities in pharmacological modelling Fundam. Clin. Pharmacol. 22 633–48
Hill A 1910 The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves J. Physiol. 40 1–7
Jakobsen J C, Gluud C, Wetterslev J and Winkel P 2017 When and how should multiple imputation be used for handling missing data in randomised clinical trials—a practical guide with flowcharts BMC Med. Res. Methodol. 17 162-017-0442-1
Jahraus A and Tobin M J 1990 Reliability of pulse oximetry in titrating supplemental oxygen therapy in ventilator-dependent patients Chest 97 1420–5
Kramer M S and Feinstein A R 1981 Clinical biostatistics. LIV. The biostatistics of concordance Clin. Pharmacol. Ther. 29 111–23
Levy M M et al 2001 SCCM/ESICM/ACCP/ATS/SIS international sepsis definitions conference Crit. Care Med. 31 1250–6
Ludbrook J 1997 Comparing methods of measurements Clin. Exp. Pharmacol. Physiol. 24 193–203
Ludbrook J 2002 Statistical techniques for comparing measurers and methods of measurement: a critical review Clin. Exp. Pharmacol. Physiol. 29 527–36
Pandharipande P P et al 2009 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. 37 1317–21
R Core Team 2018 R: a Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing) (www.R-project.org)
Rice T W et al 2007 Comparison of the SpO$_2$/FiO$_2$ ratio and the PaO$_2$/FiO$_2$ ratio in patients with acute lung injury or ARDS Chest 132 410–7
Sanz F et al 2015 Accuracy of PaO$_2$/FiO$_2$ calculated from SpO$_2$ for severity assessment in ED patients with pneumonia Respirology 20 813–8
Severinghaus J W 1979 Simple, accurate equations for human blood O$_2$ dissociation computations J. Appl. Physiol. Respir. Environ. Exerc. Physiol. 46 599–602
Seymour C W et al 2016 Assessment of clinical criteria for sepsis: for the third international consensus definitions for sepsis and septic shock (Sepsis-3) JAMA 315 762–74
Singer M et al 2016 The third international consensus definitions for sepsis and septic shock (Sepsis-3) JAMA 315 801–10
Vincent J L et al 1996 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. 22 707–10