Comparison of SpO_2 to PaO_2 based markers of lung disease severity for children with acute lung injury*

Robinder G. Khemani, MD, MsCI; Neal J. Thomas, MD, MsC; Vani Venkatachalam, MD; Jason P. Scimeme, MD; Ty Berutti, MD, MS; James B. Schneider, MD; Patrick A. Ross, MD; Douglas F. Willson, MD; Mark W. Hall, MD; Christopher J. L. Newth, MD, FRCPC; on behalf of the Pediatric Acute Lung Injury and Sepsis Network Investigators (PALISI)

Objective: **Given pulse oximetry is increasingly substituting for arterial blood gas monitoring, noninvasive surrogate markers for lung disease severity are needed to stratify pediatric risk. We sought to validate prospectively the comparability of SpO_2/FIO_2 to PaO_2/FIO_2 and oxygen saturation index to oxygenation index in children. We also sought to derive a noninvasive lung injury score.**

Design: **Prospective, multicentered observational study in six pediatric intensive care units.**

Patients: **One hundred thirty-seven mechanically ventilated children with SpO_2 80% to 97% and an indwelling arterial catheter.**

Interventions: **Simultaneous blood gas, pulse oximetry, and ventilator settings were collected. Derivation and validation data sets were generated, and linear mixed modeling was used to derive predictive equations. Model performance and fit were evaluated using the validation data set.**

Measurements and Main Results: **One thousand one hundred ninety blood gas, SpO_2, and ventilator settings from 137 patients were included. Oxygen saturation index had a strong linear association with oxygenation index in both derivation and validation data sets, given by the equation oxygen saturation index = 2.76 + 0.547*oxygenation index (derivation). $1/SpO_2/FIO_2$ had a strong linear association with $1/PaO_2/FIO_2$ in both derivation and validation data sets given by the equation $1/SpO_2/FIO_2 = 0.00232 + 0.443/PaO_2/FIO_2$ (derivation). SpO_2/FIO_2 criteria for acute respiratory distress syndrome and acute lung injury were 221 (95% confidence interval 215–226) and 264 (95% confidence interval 259–269). Multivariate models demonstrated that oxygenation index, serum pH, and $PaCO_2$ were associated with oxygen saturation index ($p < .05$); and $1/PaO_2/FIO_2$, mean airway pressure, serum pH, and $PaCO_2$ were associated with $1/SpO_2/FIO_2$ ($p < .05$). There was strong concordance between the derived noninvasive lung injury score and the original pediatric modification of lung injury score with a mean difference of −0.0361 ±0.264 SD.**

Conclusions: **Lung injury severity markers, which use SpO_2, are adequate surrogate markers for those that use PaO_2 in children with respiratory failure for SpO_2 between 80% and 97%. They should be used in clinical practice to characterize risk, to increase enrollment in clinical trials, and to determine disease prevalence. (Crit Care Med 2012; 40:1309–1316)**

KEY WORDS: acute lung injury; pediatrics; pulse oximetry; respiration, artificial; severity of illness index

Children with acute lung injury (ALI) and acute respiratory distress syndrome (ARDS) face substantial morbidity and mortality in pediatric intensive care units (PICUs) (1, 2), but there have only been a few interventional clinical trials on pediatric ALI/ARDS (3, 4). The reasons for this are multifactorial but are in part the result of the fact that the 1994 American–European Consensus Conference definition of ALI/ARDS requires an arterial blood gas (ABG) to calculate the PaO_2/FIO_2 (PF) ratio (5). Given the challenges to placing arterial catheters in small children, coupled with the increased use and reliability of pulse oximetry in PICUs (6), PaO_2 values are not routinely obtained, precluding calculation of the PF ratio. This has resulted in an inability to screen many critically ill children for ALI/ARDS (2, 6) and has resulted in an underestimation of the disease prevalence in children. In addition, ALI/ARDS studies have generally been limited to children that have more severe lung injury or hemodynamic compromise, because these are the children who receive arterial lines.

Furthermore, the optimal variable to quantify oxygenation impairment for risk

*See also p. 1386.

From the Children's Hospital Los Angeles (RGK, PAR, CJLN), University of Southern California Keck School of Medicine, Los Angeles, CA; Penn State Children's Hospital (NJT), Pennsylvania State University College of Medicine, Hershey, PA; the University of Virginia Children's Hospital (VV, DFW), University of Virginia School of Medicine, Charlottesville, VA; Nationwide Children's Hospital (JPS, MWH), The Ohio State University College of Medicine, Columbus, OH; Monroe Carell Children's Hospital (TB), Vanderbilt University School of Medicine, Nashville, TN; and Cohen Children's Medical Center of New York (JBS), North Shore Long Island Jewish Health System, New Hyde Park, NY.

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 Web site (http://journals.lww.com/ccmjournal).

Dr. Khemani received funding from the National Institutes of Health/National Library of Medicine 1RC1LM010639-01 and a fellowship from the Whittier Foundation. Dr. Willson received support from Pneuma Pharmaceuticals (50% salary support plus support for a research nurse). Dr. Thomas received support from the Food and Drug Administration 1R01FD003410-01A1. Dr. Thomas is a paid consultant and a member of the Scientific Advisory Board for Discovery Laboratories. Dr. Newth received funding from the National Institutes of Health/National Institute of Child Health and Human Development 2U10HD050012-06, National Institutes of Health/National Heart, Lung and Blood Institute 1U01HL094345-01, and a pharmacokinetics study from Hospira on dexmedetomidine in infants and children. The remaining authors have not disclosed any potential conflicts of interest.

For information regarding this article, E-mail: rkhemani@chla.usc.edu

DOI: 10.1097/CCM.0b013e31823bc61b

stratification or clinical trial enrollment remains unknown. The PF ratio is not reliably associated with mortality in adults with lung injury (7–10) (although the association is slightly stronger in children [11, 12]), and the ALI and ARDS PF cutoff values were determined by consensus, somewhat arbitrarily. An alternative, the oxygenation index (OI = 100* Paw* F_{IO_2}/PaO_2), has been used as entry criteria and to stratify risk in numerous pediatric studies (3, 4, 11–13), but also requires an ABG. OI has an inherent advantage over PF ratio in that it incorporates mean airway pressure (Paw) (14), and it may have a more reliable association than PF ratio with mortality in children (11–13). OI <6 has been used for extubation readiness testing for clinical trials (3). OI >13 has been used to describe "high-risk" patients with ARDS for clinical trials on surfactant (4). In neonates, OI >20 has been argued for consideration of high-frequency oscillatory ventilation and OI >40 for consideration of extracorporeal membrane oxygenation. Lastly, the Murray lung injury score, initially used to describe severe ARDS in adults (15), has also been used for risk stratification. It averages four components, scored 0–4: PF ratio, quadrants of alveolar consolidation on chest radiograph, dynamic compliance of the respiratory system, and level of positive end-expiratory pressure. We have previously adapted this score for pediatrics (16) and have demonstrated that it is associated with mortality in children (11). However, it too requires arterial sampling for calculation of the PF ratio.

Although it is clinically acceptable for children without lung injury to have SpO_2 between 92% and 100%, target SpO_2 for children with ALI is between 88% and 95%. Because the oxyhemoglobin dissociation curve is nearly linear when SpO_2 is between 80% and 97%, it may be possible to characterize pediatric lung injury accurately by substituting SpO_2 for PaO_2. A prospective study in adults suggested that the pulse oximetry saturation ratio (SF = SpO_2/F_{IO_2}) may be a reliable surrogate for the PF ratio when SpO_2 is ≤97% (17). We have shown similar results in retrospective studies of the SF ratio in children, although the SF ratios corresponding to PF ratios for ARDS (200) and ALI (300) were lower relative to the adult study (18, 19). SF ratios have also been associated with mortality in children (20). Our retrospective data also demonstrate that oxygen saturation index [OSI = 100*Paw*F_{IO_2}/SpO_2], which substitutes SpO_2 for PaO_2, adequately discriminates PF criteria for ALI and ARDS (19) and is associated with mortality (20).

The goal of this study was to test whether noninvasive, SpO_2-based measures of pediatric oxygenation impairment can be used to predict values that, to date, have only been obtainable through ABG sampling. We therefore sought to prospectively validate the comparability of OSI to OI and SF to PF in a multicenter observational study in children. Furthermore, we sought to derive and validate a noninvasive version of the pediatric lung injury score (NLIS). We reasoned validation of these noninvasive markers will expand the population of children eligible for clinical studies on ALI/ARDS, offer alternative diagnostic criteria, and facilitate gathering accurate epidemiologic data related to pediatric ALI and ARDS.

MATERIALS AND METHODS

Patients. We conducted a prospective, observational, cohort study in six multidisciplinary U.S. PICUs from August 2009 to October 2010. Participating PICUs included: Children's Hospital Los Angeles, Penn State Children's Hospital, the University of Virginia Children's Hospital, Nationwide Children's Hospital, Monroe Carell Children's Hospital, and Cohen Children's Medical Center of New York. Children (≥37 wks gestational age and <18 yrs) were included if they were intubated and mechanically ventilated with an indwelling arterial catheter and had continuous pulse oximeter recordings. Measurements from children were not recorded if there was an inadequate pulse oximetry waveform, SpO_2 measurements were >97%, the patient was on extracorporeal membrane oxygenation, had cyanotic congenital heart disease, or was not on a fully supported mode of ventilation (i.e., pressure support or continuous positive airway pressure alone). Measurements in which SpO_2 was <80% were gathered but excluded from analysis. The study was approved by the institutional review boards at all participating hospitals with a waiver of written consent.

Study Measurements. Obtaining an ABG sample was left to the discretion of the primary care team. SpO_2 values were documented at the precise times the ABG samples were drawn. To ensure accurate results, study personnel checked the position and cleanliness of the pulse oximetry sensor. ABG and SpO_2 pairs were not recorded on the data collection sheet if the patient had received endotracheal suctioning or position changes in the 10 mins before the ABG or ventilator changes or invasive procedures in the 30 mins before measurement. Practitioners at all participating PICUs titrated F_{IO_2} to maintain oxygen saturation between 88% and 97%. Three institutions used Nellcor pulse oximeters (Covidien-Nellcor, Boulder, CO), and three used Masimo pulse oximeters (Masimo Corporation, Irvine, CA).

On enrollment in the study, we collected diagnosis, age, race, gender, and weight for each patient. With each eligible ABG, we recorded data regarding pulse oximetry, the results of the ABG, the degree of ventilator support, capillary refill in the extremity where the pulse oximeter probe was located, end-tidal CO2, temperature, and hemoglobin. Hemoglobin values were recorded from laboratory samples and, if unavailable, were recorded from the blood gas. We reviewed the most recent chest radiograph (before the ABG) for the presence of bilateral pulmonary infiltrates and number of quadrants of alveolar consolidation.

Analysis. The primary objective was to determine the association between OSI and OI. Secondary objectives included the strength of association between SF ratio and PF ratio as well as the derivation of a noninvasive counterpart for the pediatric modification of the Murray lung injury score (NLIS) (15, 16). Analytic methods are primarily described in relation to OSI and OI. Methods were nearly identical for the relationship between SF and PF. Methods for NLIS are described at the end of this section. Statistical analysis was performed in Statistica version 9 (Statsoft, Tulsa, OK) and Stata version 10 (StatCorp, College Station, TX).

Derivation Data Set. The data set was divided into derivation and validation groups through computer-generated random number assignment with a 60%:40% split. Descriptive statistics and two-way scatterplots characterized the relationship between OI and OSI. Given repeated measurements per patient, we used linear mixed modeling to derive a predictive equation for the relationship between OSI and OI to control for patient-specific clustering. Patient number was introduced as the random effect in the model with OI as the fixed effect. To evaluate the benefit of adding the fixed effects to the random effects model alone, we report Akaike Information Criteria, Schwarz' Bayesian criterion, and log likelihood. Lower values of Akaike Information Criteria and Schwarz' Bayesian criterion indicate better fit and model parsimony.

Based on the derived regression equation, we determined OSI values that correspond to clinically important OI values (6, 13, 20, and 40). We calculated the sensitivity, specificity, false-positive rate, false-negative rate, and positive and negative likelihood ratios of the OSI cutoffs against the gold standard OI cutoffs. Area under the receiver operating characteristic plots evaluated the discrimination ability of OSI on each of the clinically important OI cutoffs.

Validation Data Set. We used the validation dataset for two purposes. First, we used the validation data set to generate predictive equations and corresponding cutoff values for the relationship between OI and OSI to demonstrate reproducibility between validation and derivation data sets. Second we sought to apply the predicted cutoffs for OSI from the

Table 1. Descriptive statistics for individual observations (blood gas and pulse oximetry pairs) stratified by derivation and validation data sets

	All (n = 1190)	Derivation (n = 721)	Validation (n = 469)
SpO_2, %	95 (92–96)	94 (92–96)	95 (92–96)
PaO_2, mm Hg	70 (61–83)	70 (62–84)	69 (61–82)
$Paco_2$, mm Hg	54 (45–67)	53 (45–66)	55 (45–68)
pH	7.40 (7.33–7.45)	7.40 (7.33–7.45)	7.40 (7.32–7.45)
FIO_2	0.56 (0.40–0.65)	0.56 (0.40–0.65)	0.55 (0.40–0.65)
Mean airway pressure, cm H_2O, all modes	18 (13–24)	18 (13–24)	18 (13–24)
Temperature, °C	37.0 (3.7.0–37.0)	37.0 (37.0–38.0)	37.0 (37.0–38.0)
Hemoglobin, g/dL	11 (10–13)	12 (10–13)	11 (10–13)
Capillary refill, secs	<2 (<2–3)	<2 (<2–3)	<2 (<2–3)
Bilateral infiltrates, no. (%)	825 (69.3)	495 (68.7)	330 (70.4)
Quadrants of consolidation	4 (2–4)	4 (2–4)	4 (2–4)
Conventional ventilation	(n = 916)	(n = 572)	(n = 344)
Positive end-expiratory pressure, cm H_2O	8 (5–10)	8 (5–10)	8 (5–10)
Tidal volume, mL/kg	8 (7–11)	8 (7–11)	8 (7–11)
Peak inspiratory pressure, cm H_2O	29 (25– 35)	29 (25– 35)	30 (25–36)
Pressure support, cm H_2O	10 (5–12)	10 (5–12)	10 (5–13)
Ventilator rate, breaths/min	20 (15–24)	20 (16–24)	20 (15–24)
Dynamic compliance of the respiratory system, mL/cm H_2O/kg	0.40 (0.31–0.57)	0.40 (0.32–0.58)	0.40 (0.31–0.55)
High-frequency oscillatory ventilation	(n = 274)	(n = 149)	(n = 125)
Mean airway pressure, cm H_2O	30 (26–34)	30 (25–33)	30 (26–34)
Hertz	8 (6–9)	8 (6–9)	8 (6–10)
Amplitude	58 (48–65)	58 (46–62)	57 (50–70)
Oxygenation index	13.4 (7.4–23.4)	13.4 (7.3–22.6)	13.3 (7.4–24.8)
Oxygen saturation index	10.2 (5.9–17.4)	10.3 (6.0–17.2)	10.1 (5.9–18.0)
PaO_2/FIO_2	133 (93–187)	133 (95–192)	133 (92–176)
SpO_2/FIO_2	164 (139–233)	164 (140–233)	162 (139–233)
Lung injury score	3.0 (2.0–3.5)	3.0 (2.0–3.5)	3.0 (2.25–3.5)
Noninvasive lung injury score	3.0 (2.0–3.5)	2.75 (2.0–3.5)	3.0 (2.0–3.5)

Data presented as median (interquartile range) or count (percentage).

derivation data set to the validation data set to determine the calibration of the model. We classified each observation in the validation data set as falling above or below each OSI cutoff from the derivation data set. We calculated the sensitivity, specificity, false-positive rate, false-negative rate, and positive and negative likelihood ratios of the OSI cutoffs from the derivation data set against the gold standard OI cutoffs.

Sensitivity Analysis. We performed sensitivity analyses to evaluate if the way we split our data sets (60%/40%) introduced bias into our group assignment. We randomly divided the entire data set into five different split samples (50%/50%, 60%/40%, 70%/30%, 80%/20%, 90%/10%) and repeated our analysis. We created linear mixed models as described previously and compared corresponding cutoffs for OSI for each of the split derivation and validation data sets.

SF Ratio. We analyzed the secondary outcome evaluating the relationship between SF ratio and PF ratio in a similar fashion as OSI and OI. We transformed data (1/PF and 1/SF) to satisfy assumptions of normality and improve model fit. We attempted other transformations (square root and other polynomials), but they were not superior to 1/PF and 1/SF.

Multivariate Analysis. We used multivariate regression using linear mixed modeling to determine whether other elements of ventilator support, race, age, temperature, pH, capillary refill, or hemoglobin confounded or modified the relationship between OSI and OI. All variables had biological plausibility for modifying or confounding the relationship between PaO_2 and SpO_2. We considered variables for model inclusion if they had a univariate relationship ($p < .2$) with either OSI or OI (or SF and PF). We kept variables in the final multivariate model if they confounded the relationship between OSI and OI or SF and PF by changing the parameter estimates by at least 20% or if they remained statistically significantly associated with the outcome variable after controlling for other factors. To evaluate the benefit of adding additional variables to the multivariate models, we again report Akaike Information Criteria, Schwarz' Bayesian criterion, and log likelihood. We examined statistical significance of improvement in model fit by using a chi-squared test on the difference between log likelihood in the two models (21).

Noninvasive Lung Injury Score. For the development of the NLIS, we generated SF cutoff values from the derivation data set corresponding to PF breakpoints in the lung injury score (300, 225, 175, 100). These SF cut points were combined with the other elements of the lung injury score to create a NLIS. We used the entire data set and evaluated for bias using Bland-Altman plots controlling for multiple observations per patient (22).

RESULTS

There were 1207 simultaneous ABG, pulse oximetry, and ventilator settings recorded. We excluded 17 observations in which the SpO_2 <80%, leaving 1190 values from 137 patients for analysis. The observations were randomly divided (60%/40%) into derivation (n = 721) and validation (n = 469) data sets. Nearly all observations met oxygenation criteria for ALI (95.0%), and close to 80% met oxygenation criteria for ARDS. The median PF ratio was 133 with a median OI of 13.4. Bilateral infiltrates were present in 69% of chest radiographs with a median of four quadrants of alveolar consolidation. These findings were nearly identical in derivation and validation sets (Table 1). Clinicians most commonly used pressure-regulated volume control (42.7%) followed by pressure control (29.7%) and high-frequency oscillatory ventilation (23.5%). Volume control was used very infrequently (4%). The median age of patients was 3 yrs (interquartile range, 0.42–11.2) with 61.3% being male, 51% white, 30% Latino, 10.2% black, 2.9% Asian, and 5.9% other.

Derivation Data Set. OSI had a strong linear association with OI (Fig. 1A) given by the regression equation OSI = 2.76 + 0.547*OI. Clinically important cutoff values for OI (6, 13, 20, and 40) had predictable corresponding values of OSI with narrow confidence intervals (Table 2). The predicted OSI values (6.0, 9.9, 13.7, 24.7) generated positive likelihood ratios that were all >7 and negative likelihood ratios that were all <0.2, indicating OSI values have a moderate to large impact on ruling in whether the patient is above the given OI cutoff (Table 3). The overall ability of OSI to discriminate clinically important cutoff values for OI was outstanding (23) with area under the receiver operating characteristic curve plots all >0.95 (Table 3).

The relationship between SF ratio and PF ratio was best expressed using 1/SF ratio and 1/PF ratio (Fig. 1B). The linear relationship was strong given by the regression equation 1/SF = 0.00232 + 0.443/PF. The ARDS PF value of 200 corresponds to an SF value of 221 (95% confidence interval 215–226) and the ALI PF value of 300 to an SF value of 264 (95% confidence interval 259–269) (Table 2). The SF ALI value of 264 had 91% sensitivity and 53% specificity in detecting cases in which the PF ratio was <300 and the SF ARDS value of 221 had 88% sensitivity

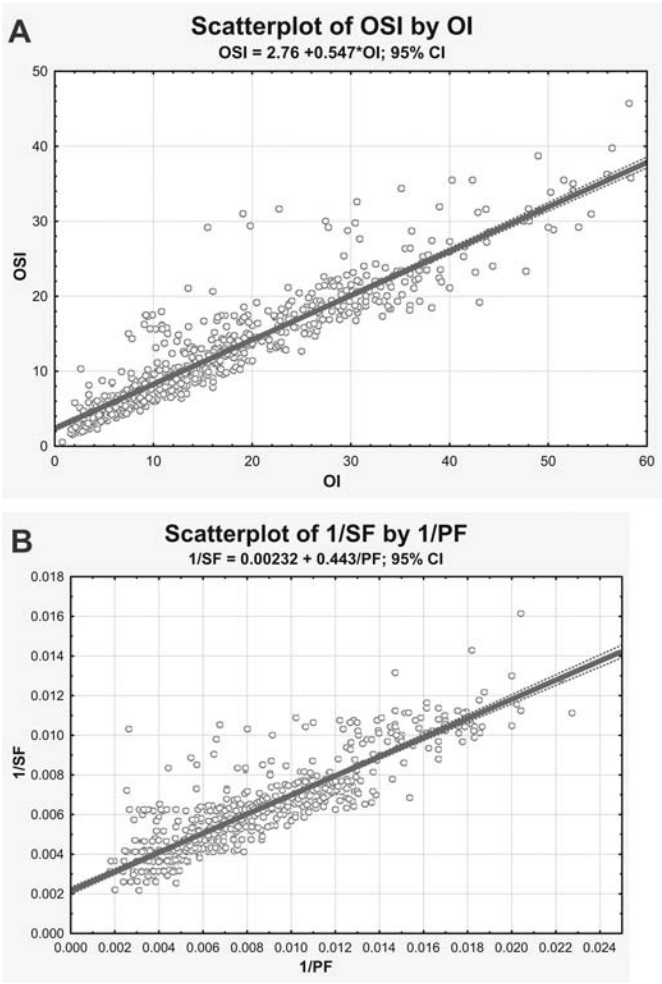

Figure 1. Scatterplot oxygen saturation index (*OSI*) vs. oxygenation index (*OI*) (*top, A*) and of $1/SpO_2/FIO_2$ (*SF*) ratio vs. $1/PaO_2/FIO_2$ (*PF*) ratio (*bottom, B*). Note the strong linear fit across the entire range of values. *CI*, confidence interval.

and 78% specificity in detecting cases in which the PF ratio was <200 (Table 3). The SF negative likelihood ratios for oxygenation criteria for ALI and ARDS were both <0.2 and positive likelihood ratios for oxygenation criteria for ALI and ARDS were 1.93 and 4.0, respectively (Table 3). Overall SF ratio had excellent (23) discrimination ability for PF values of 200 (area under the receiver operating characteristic curve 0.90) and 300 (area under the receiver operating characteristic curve 0.82) (Table 3).

Validation Data Set. The regression equation, OSI cutoffs (Table 2), and discrimination ability (Table 3) of OSI on OI cut points from the predictive model generated from the validation data set were nearly identical to the derivation data set. When the predicted OSI values from the derivation data set were applied to the validation data set, there was nearly identical sensitivity, specificity, and positive and negative likelihood ratios as the derivation data set (Table 3).

The regression equation, SF values for oxygenation criteria for ALI and ARDS (Table 2), and discrimination ability (Table 3) of SF on PF ratios of 200 and 300 from the validation data set were similar to the derivation data set. The SF ALI value of 264 and ARDS value of 221 from the derivation data set had similar sensitivity, specificity, and positive and negative likelihood ratios when applied to the validation data set (Table 3).

We performed sensitivity analysis using five different randomly generated split derivation/validation samples. Parameter estimates for OSI corresponding to clinically relevant values of OI and parameter estimates for SF corresponding to PF ratios of 200 and 300 were nearly identical in all five iterations of derivation/validation data sets (analysis not shown). This demonstrates that the generated cutoff values for OSI and SF ratio are not substantially influenced by the size of derivation/validation samples.

Multivariate Modeling. Multivariate models on the derivation data set using linear mixed models with patient number as the random effect demonstrated that OI, serum pH, and $Paco_2$ were associated with OSI ($p < .05$); and 1/PF ratio, Paw, pH, and $Paco_2$ were associated with 1/SF ratio ($p < .05$) (see Table ESM 1 [Supplemental Digital Content 1, http://links.lww.com/CCM/A364]). A 0.10 fall in pH was associated with an increase of OSI by 0.4 and a 20-mm Hg rise in $Paco_2$ with an increase of OSI by 1.1. A 0.10 fall in pH was associated with a decrease of SF by 7.0 and a 20-mm Hg rise in $Paco_2$ with a decrease of SF by 6.6. A 5-point rise in Paw was associated with a decrease of SF by 9.9. The iterative addition of these covariates to the random effects model demonstrated significant improvements in model fit with reductions in Akaike Information Criteria and Schwarz' Bayesian criterion (see Table ESM 2 [Supplemental Digital Content 1, http://links.lww.com/CCM/A364]). Because of the possible effects of fetal hemoglobin, we modeled age three ways: as a continuous variable, stratified according to neonatal age (≤1 month), or possible persistence of fetal hemoglobin (≤6 months). None of these age categories were significant in the previous multivariate models ($p > .2$). Capillary refill on the extremity where the pulse oximeter was located, weight, temperature, hemoglobin, gender, race, and quadrants of alveolar consolidation on chest radiograph were also not significantly associated with OSI or 1/SF ($p > .1$).

NLIS Derivation. We used the generated predictive equation for 1/SF ratio from 1/PF ratio from the derivation data set to determine SF cutoff points for the four ranges of Lung injury in the modified Murray lung injury score. Computation of the scores is detailed in Table 4. For the entire data set, overall the concordance was excellent with a median difference between NLIS and Murray lung injury score of 0 (interquartile range, −0.25 to 0) and a mean of -0.0361 ± 0.264 SD. Given repeated measures per patient, Bland-Altman analysis controlling for both between-subject and within-subject variance (Fig. 2) demonstrated that 1119 of 1190 observations (94%) fall within the generated 95% confidence interval of −0.55 to 0.48. However, 69 of the 71 observations

Table 2. Derived regression equations from the derivation and validation datasets from the mixed model[a]

	Derivation Set	
	Equation OSI = 2.76 + 0.547*OI	
OI	OSI (95%CI)	OSI (99% CI)
≥6	6.0 (5.9–6.2)	6.0 (5.9–6.2)
≥13	9.9 (9.6–10.2)	9.9 (9.5–10.2)
≥20	13.7 (13.3–14.1)	13.7 (13.2–14.3)
≥40	24.7 (23.8–25.5)	24.7 (23.5–25.8)
	Equation 1/SF = 0.00232 + 0.443/PF	
PF	SF (95% CI)	SF (99% CI)
≤200	221 (215–226)	221 (214– 228)
≤300	264 (259–269)	264 (257– 270)
	Validation Set	
	Equation OSI = 2.55 + 0.570*OI	
OI	OSI (95% CI)	OSI (99% CI)
≥6	6.0 (5.8–6.1)	6.0 (5.8–6.2)
≥13	10.0 (9.7–10.3)	10.0 (9.6–10.4)
≥20	14.0 (13.5–14.4)	14.0 (13.3–14.6)
≥40	25.4 (24.4–26.3)	25.4 (24.1–26.6)
	Equation 1/SF = 0.00226 + 0.461/PF	
PF	SF (95% CI)	SF (99% CI)
≤200	219 (213–226)	219 (211–228)
≤300	263 (257–270)	263 (255–272)

OSI, oxygen saturation index; OI, oxygenation index; CI, confidence interval; SF, SpO_2/FIO_2; PF, PaO_2/FIO_2.

[a]To the right are the corresponding OSI and SF values for clinically relevant OI and PF values with both 95% and 99% CI.

that fell outside of this range were above the upper confidence interval. Graphically, it appears as if there is similar variability across all values of lung injury score.

DISCUSSION

This study demonstrates that in children with acute hypoxemic respiratory failure, lung injury severity markers that use SpO_2 are adequate surrogate markers for those that use PaO_2 as long as SpO_2 is between 80% and 97%. Given the now routine use of pulse oximetry in nearly all intensive care units, and the more infrequent use of arterial lines in children, noninvasive oxygenation criteria should be strongly considered to characterize patient risk and as enrollment criteria for clinical trials in children when a PaO_2 is unavailable.

The ALI and ARDS SF cutoff values of 264 and 221, respectively, in this study are similar to previously reported values in children (18, 19) but lower than those reported in adult studies (17, 24). The potential reasons are multiple. Higher values of SpO_2 approach the nonlinear portion of the oxyhemoglobin dissociation curve. Although these effects were minimized by transforming the data, close to 40% of SpO_2 values were 96% or 97%. The model fit improves slightly if SpO_2 is 80% to 95% (analysis not shown) with slightly higher SF cutoffs (275 and 226). However, it was difficult to convince providers to maintain oxygen saturations ≤97% consistently and will be even more difficult to target saturations ≤95%. Nonetheless, it is clear that the accuracy of these measurements can be improved by targeting FIO_2 to achieve SpO_2 between 88% and 95%. This goal may also help minimize the potential effects of oxygen toxicity (25).

Although fetal hemoglobin can persist for the first 3–6 months of life, age was not significant in any of the multivariate models, and when restricted to children <6 months, parameter estimates of SF ratio for oxygenation criteria for ALI and ARDS were nearly identical to the entire population. It does not appear that fetal hemoglobin explains the lower SF cutoff values seen in children; in fact, fetal hemoglobin should shift the oxyhemoglobin dissociation curve to the left, yielding a relatively higher SF ratio for a given PF ratio.

The relationship between SpO_2 and PaO_2 can be altered by pH, $Paco_2$, temperature, and concentration of 2,3-diphosphoglycerate. The multivariate models demonstrate that pH, $Paco_2$, and Paw have an effect on the relationship between SpO_2 and PaO_2. Children in this study had similar average pH, lower PF ratios, slightly lower positive end-expiratory pressure (for conventional ventilation), and higher $Paco_2$ (54 mm Hg vs. 41 mm Hg) compared with adults in the study by Rice et al (17). Accepting the frequent use of high-frequency oscillatory ventilation, we accounted for ventilator effect using Paw rather than positive end-expiratory pressure (17, 24). From our multivariate model, a 5-point increase in Paw results in a 10-point fall in SF ratio. The reason why higher Paw is associated with a lower SF ratio for a given PF ratio is not entirely clear but was also seen in the study by Rice (with positive end-expiratory pressure) (17). This may be secondary to cardiopulmonary interactions with higher degrees of ventilator support. The higher $Paco_2$ noted in this study compared with the adult study would correspond to a negligible 4-point decrease in the SF cutoff values for oxygenation criteria for ALI and ARDS assuming that Paw is unchanged. Unfortunately, levels of 2,3-diphosphoglycerate were not measured. Although the effects of pH and $Paco_2$ were statistically significant in the model predicting OSI, the parameter estimates were small and their impact on clinical practice is likely insignificant. This further supports the notion that some measure of ventilator support (like in OI, OSI, lung injury score, NLIS) should be included in lung disease severity markers, because the degree of ventilator support appears to affect the relationship between SF and PF ratio for both children and adults with lung injury.

Furthermore, our cohort of children had acute hypoxemic respiratory failure, not necessarily ALI. Just over two-thirds

Table 3. Discrimination ability of the $SpO_2/F[scap]io[r]_2$ ratio against $PaO_2/F[scap]io[r]_2$ criteria for acute lung injury and acute respiratory distress syndrome as well as the discrimination of oxygen saturation index against clinically significant oxygenation index values from the derivation and validation data sets[a]

	Derivation	Validation[b]
Oxygenation criteria for acute lung injury ($PaO_2/FIO_2 \leq 300$)		
Area under the receiver operating characteristic curve	0.82	0.93
Sensitivity (true-positive)	0.91	0.92
Specificity (true-negative)	0.53	0.64
False-positive rate	0.47	0.36
False-negative rate	0.09	0.08
Positive likelihood ratio	1.93	2.52
Negative likelihood ratio	0.17	0.13
No. (%) of cases classified correctly	643 (89%)	424 (90%)
Oxygenation criteria for acute respiratory distress syndrome ($PaO_2/FIO_2 \leq 200$)		
Area under the receiver operating characteristic curve	0.90	0.92
Sensitivity (true-positive)	0.88	0.89
Specificity (true-negative)	0.78	0.80
False-positive rate	0.22	0.20
False-negative rate	0.12	0.11
Positive likelihood ratio	4.00	4.40
Negative likelihood ratio	0.16	0.14
No. (%) of cases classified correctly	618 (86%)	408 (87%)
Oxygenation index ≥ 6		
Area under the receiver operating characteristic curve	0.96	0.97
Sensitivity (true-positive)	0.87	0.91
Specificity (true-negative)	0.92	0.95
False-positive rate	0.08	0.05
False-negative rate	0.13	0.09
Positive likelihood ratio	10.40	17.4
Negative likelihood ratio	0.15	0.10
No. (%) of cases classified correctly	631 (88%)	428 (91%)
Oxygenation index ≥ 13		
Area under the receiver operating characteristic curve	0.97	0.98
Sensitivity (true-positive)	0.92	0.94
Specificity (true-negative)	0.91	0.91
False-positive rate	0.09	0.09
False-negative rate	0.08	0.06
Positive likelihood ratio	9.74	10.93
Negative likelihood ratio	0.09	0.06
No. (%) of cases classified correctly	658 (91%)	436 (93%)
Oxygenation index ≥ 20		
Area under the receiver operating characteristic curve	0.98	0.97
Sensitivity (true-positive)	0.94	0.96
Specificity (true-negative)	0.89	0.88
False-positive rate	0.11	0.12
False-negative rate	0.06	0.05
Positive likelihood ratio	8.40	7.64
Negative likelihood ratio	0.06	0.05
No. (%) of cases classified correctly	652 (90%)	423 (90%)
Oxygenation index ≥ 40		
Area under the receiver operating characteristic curve	0.99	0.98
Sensitivity (true-positive)	0.88	0.88
Specificity (true-negative)	0.97	0.95
False-positive rate	0.03	0.05
False-negative rate	0.12	0.12
Positive likelihood ratio	29.85	17.88
Negative likelihood ratio	0.13	0.13
No. (%) of cases classified correctly	696 (97%)	443 (94%)

[a]Data presented as area under the receiver operating characteristic curve plots. These values are generated from separate predictive models in each data set. Sensitivity, specificity, false-positive rate, false-negative rate, positive and negative likelihood ratios, and number and percent of correctly classified cases using the derived SpO_2/FIO_2 ratio equivalents for acute lung injury (264) and acute respiratory distress syndrome (221) from the derivation data set against the PaO_2/FIO_2 criteria for acute lung injury (300) and acute respiratory distress syndrome (200) applied to both derivation and validation data sets. Sensitivity, specificity, false-positive rate, false-negative rate, positive and negative likelihood ratios, and number and percent of correctly classified cases using the derived oxygen saturation index equivalents (6.0, 9.9, 13.7, 24.7) for clinically important oxygenation index cutoffs from the derivation data set applied to both derivation and validation data sets; [b]cutoff values generated from the derivation data set applied to the validation data set.

of observations (blood gas/pulse oximetry pairs) had a corresponding chest radiograph that demonstrated bilateral pulmonary infiltrates, and we did not specifically track left ventricular dysfunction. Therefore, a true estimate of the prevalence of ALI/ARDS is not possible from our data but was not an objective. In contrast, all patients in the Rice study met all four ALI criteria. Although there may be unmeasured variables that alter the relationship between SF and PF ratios seen only in patients with all four ALI criteria, this is unlikely because chest radiograph findings were not significant in our multivariate model. Furthermore, a subsequent study in adults without ALI has also demonstrated SF cutoff values for PF ratios of 200 and 300, which are higher (24) than our reported values in children.

The specificity of SF criteria for PF ratios of 300 was low (53%), which may lead some to conclude that it may not be a useful surrogate for PF ratio. However, it is important to remember that measures of sensitivity and specificity are dependent on the underlying prevalence of disease, a concept that is highly relevant to our cohort given than only 5% of our subjects did not meet oxygenation criteria for ALI. Likelihood ratios, on the other hand, generate a clinically interpretable posttest probability that the patient actually has the disease, wherein a likelihood ratio >1 suggests the test result is associated with disease and a likelihood ratio (LR) <1 suggests the test result is associated with absence of disease. In our population, the prevalence of oxygenation criteria for ALI was very high (approximately 95%). With the positive LR of 2, the posttest probability that a patient will have a PF ratio <300 if their SF ratio <264 is 97.5%. The negative LR of 0.17 yields a posttest probability of 76%. The prevalence of oxygenation criteria for ARDS was approximately 80%, and the positive LR of 4 yields a posttest probability of 94%, and the negative LR of 0.16 yields a posttest probability of 39%. Therefore, in this population, patients who have SF ratios <264 or <221 have a very high probability of meeting PF criteria for ALI or ARDS, respectively. However, an SF ratio >221 or 264 does not reliably rule out the presence of oxygenation criteria for ARDS or ALI, respectively.

When generalizing these results to other populations, it is important to remember that SF criteria should only be used for mechanically ventilated patients with SpO_2 between 80% and 97%.

Table 4. Original pediatric modified lung injury score, and the newly derived noninvasive equivalent substituting SpO_2/FIO_2 ratio for PaO_2/FIO_2 ratio[a]

	Score				
	0	1	2	3	4
Modified pediatric lung injury score					
Quadrants of consolidation	0	1	2	3	4
PaO_2/FIO_2 ratio	≥300	225–300	175–225	100–175	≤100
Positive end-expiratory pressure, cm H_2O	≤4	5–6	7–8	9–11	≥12
Dynamic compliance of the respiratory system, mL/cm H_2O/kg	>0.85	0.75–0.85	0.55–0.74	0.30–0.54	<0.3
Modified pediatric noninvasive lung injury score					
Quadrants of consolidation	0	1	2	3	4
SpO_2/FIO_2 ratio	≥264	233–264	206–233	148–206	≤148
Positive end-expiratory pressure, cm H_2O	≤4	5–6	7–8	9–11	≥12
Dynamic compliance of the respiratory system, mL/cm H_2O/kg	>0.85	0.75–0.85	0.55–0.74	0.30–0.54	<0.3

[a]All four elements are scored 0–4, and the average of the four components is taken to generate the score. If an element is missing (e.g., compliance, positive end-expiratory pressure for high-frequency oscillatory ventilation), then the average of available components is taken.

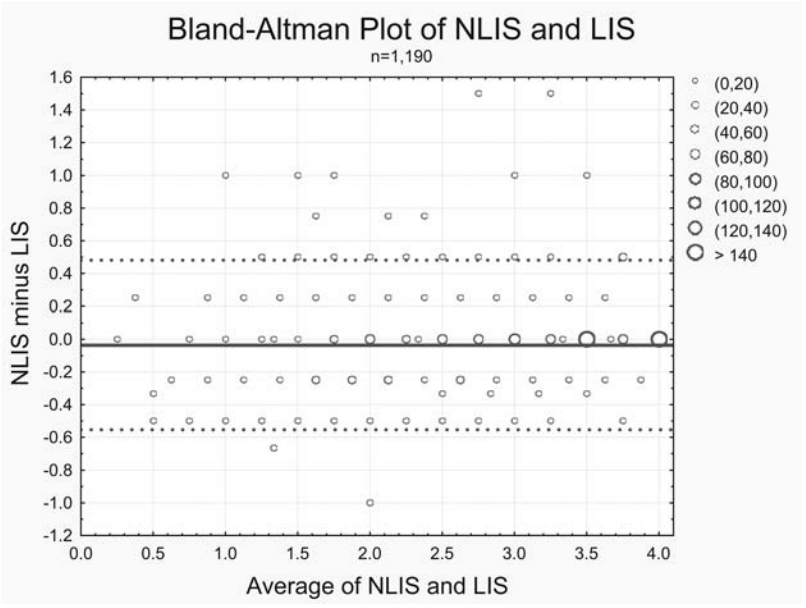

Figure 2. Bland-Altman plot of lung injury score (*LIS*) and the noninvasive LIS (*NLIS*) adjusted for repeated measurements. The y-axis subtracts the LIS from the NLIS, and the x-axis is the average of the two scores. The size of each point on the graph represents a range in number of observations corresponding to the scale to the right of the graph. The *solid line* is the mean difference with the *dotted lines* representing the 95% confidence interval of the mean differences. There appears to be no significant systematic bias with the mean difference between the LIS and the NLIS very close to zero. There appears to be equal variability across the entire range of average scores indicating no proportional error. There are more values which fall above the upper 95% confidence interval than below the lower 95% confidence interval, indicating that although there is no systematic error, NLIS may be more likely to overestimate LIS than underestimate LIS.

These populations will likely be similar to our study cohort with a high prevalence of patients who will also have PF ratios <300. Furthermore, if FIO_2 can be weaned to target SpO_2 between 88% and 95%, then not only will the potential accuracy of the SF ratio increase, but the prevalence of patients meeting oxygenation criteria for ALI will remain high. This enhances the use of the SF ratio to accurately rule in oxygenation criteria for ALI or ARDS, facilitating its use as inclusion criteria for clinical trials because one will be very confident (97.5%, assuming the same prevalence) that the patient would meet PF criteria for ALI or ARDS if they meet SF criteria.

OSI values that correspond to OI values that have been used in clinical trials (≥6, ≥13) have excellent positive and negative LRs. Therefore, if a patient meets certain OSI criteria, they have a very high likelihood of meeting the corresponding OI criteria. The OSI value of 9.9 corresponding to the OI value of 13 has a positive LR close to 10 and negative LR <0.1, lending its use to studies involving therapies targeting patients with more significant lung disease, as previously used in the calfactant trial (4).

We have also derived, for the first time, a noninvasive version of the Murray lung injury score, which has previously been modified for pediatrics (16). The derived score is nearly identical to the invasive score across all ranges of lung disease severity. The SF-derived scores for NLIS may have more variability when SF ratios grade higher lung disease severity than PF ratios (as seen by the number of observations above the 95% CI), but this difference is small and likely clinically insignificant. Because the lung injury score is a composite measure of lung disease severity, it may be a more robust metric to stratify risk for clinical trials and quality improvement. Validating a noninvasive version of the score will allow it to be applied to a broader range of patients.

Although we have demonstrated that SF criteria can substitute for PF criteria to aid in the diagnosis of ALI or ARDS, it is also clear that this relationship is confounded by degree of ventilator support. Given the less frequent use of arterial lines, the somewhat arbitrary nature of the PF ratio values generated from the consensus guidelines, the inconsistent relationship between PF ratio and outcome for both adults and children with lung injury, and the impact of ventilator support on the relationship between PF and SF ratios, perhaps it is time to reconsider the diagnostic criteria for ALI and ARDS. We have provided two other candidate lung injury severity markers that are consistent when using either SpO_2 or PaO_2 (OSI and OI, NLIS and Murray lung injury score).

There are some limitations to our study. Although we attempted to gather information on a multitude of potential confounding variables, some variables had no reliable surrogate marker (e.g., 2,3-diphosphoglycerate), and there may be

important confounders we did not consider. In addition, although capillary refill was not a significant confounder in our multivariate models, poor perfusion states may decrease the reliability of pulse oximetry measurements. We specifically excluded patients with unreliable tracings, but when applying these diagnostic criteria in other situations, it is important to only calculate these metrics when there is a reliable pulse oximetry value. Finally, we could not assess for selection bias because we did not track eligible but not enrolled patients. However, given the nature of this study, that likely has a minimal impact on the results.

CONCLUSIONS

We should strongly consider using noninvasive markers of oxygenation to characterize risk, to increase enrollment in clinical trials, and to gather accurate epidemiologic data related to pediatric lung injury. We have provided prospective validation of three such metrics in children and support their use in clinical practice and for future investigations in children with lung injury when an ABG is not available. SF ratios of 221 and 264 correspond to PF ratios of 200 and 300, and OSI values of 6.0, 9.9, 13.7, and 24.7 correspond to OI values of 6, 13, 20, and 40, respectively. NLIS scores are equivalent to their invasive counterpart. Furthermore, the accuracy of these measurements can be improved if targets of SpO_2 are consistently between 88% and 95%, a notion that has been embraced by our neonatology colleagues to limit oxygen toxicity.

ACKNOWLEDGMENTS

The authors thank the research coordinators: Debbie Spear, Jill Raymond, Christine Traul and Jeff Terry; their programmer, Paul Vee; and all the critical care nurses and respiratory therapists from the six PICUs for their help.

REFERENCES

1. Erickson S, Schibler A, Numa A, et al: Acute lung injury in pediatric intensive care in Australia and New Zealand: A prospective, multicenter, observational study. *Pediatr Crit Care Med* 2007, 8:317–323
2. Santschi M, Jouvet P, Leclerc F, et al: Acute lung injury in children: Therapeutic practice and feasibility of international clinical trials. *Pediatr Crit Care Med* 2010; 11:681–689
3. Curley MA, Hibberd PL, Fineman LD, et al: Effect of prone positioning on clinical outcomes in children with acute lung injury: A randomized controlled trial. *JAMA* 2005; 294:229–237
4. Willson DF, Thomas NJ, Markovitz BP, et al: Effect of exogenous surfactant (calfactant) in pediatric acute lung injury: A randomized controlled trial. *JAMA* 2005; 293:470–476
5. Bernard GR, Artigas A, Brigham KL, et al: The American–European Consensus Conference on ARDS. Definitions, mechanisms, relevant outcomes, and clinical trial coordination. Am *J Respir Crit Care Med* 1994; 149:818–824
6. Khemani RG, Markovitz BP, Curley MA: Characteristics of children intubated and mechanically ventilated in 16 PICUs. *Chest* 2009; 136:765–771
7. Montgomery AB, Stager MA, Carrico CJ, et al: Causes of mortality in patients with the adult respiratory distress syndrome. *Am Rev Respir Dis* 1985; 132:485–489
8. Zilberberg MD, Epstein SK: Acute lung injury in the medical ICU: Comorbid conditions, age, etiology, and hospital outcome. *Am J Respir Crit Care Med* 1998; 157:1159–1164
9. Nuckton TJ, Alonso JA, Kallet RH, et al: Pulmonary dead-space fraction as a risk factor for death in the acute respiratory distress syndrome. *N Engl J Med* 2002; 346:1281–1286
10. Doyle RL, Szaflarski N, Modin GW, et al: Identification of patients with acute lung injury. Predictors of mortality. *Am J Respir Crit Care Med* 1995; 152:1818–1824
11. Khemani RG, Conti D, Alonzo TA, et al: Effect of tidal volume in children with acute hypoxemic respiratory failure. *Intensive Care Med* 2009; 35:1428–1437
12. Flori HR, Glidden DV, Rutherford GW, et al: Pediatric acute lung injury: Prospective evaluation of risk factors associated with mortality. *Am J Respir Crit Care Med* 2005; 171:995–1001
13. Trachsel D, McCrindle BW, Nakagawa S, et al: Oxygenation index predicts outcome in children with acute hypoxemic respiratory failure. *Am J Respir Crit Care Med* 2005; 172:206–211
14. Chan KPW, Stewart TE: Clinical use of high-frequency oscillatory ventilation in adult patients with acute respiratory distress syndrome. *Crit Care Med* 2005; 33(Suppl):S170–S174
15. Murray JF, Matthay MA, Luce JM, et al: An expanded definition of the adult respiratory distress syndrome. *Am Rev Respir Dis* 1988; 138:720–723
16. Hammer J, Numa A, Newth CJ: Acute respiratory distress syndrome caused by respiratory syncytial virus. *Pediatr Pulmonol* 1997; 23:176–183
17. Rice TW, Wheeler AP, Bernard GR, et al; for the 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
18. Khemani RG, Patel NR, Bart RD, 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
19. Thomas N, Shaffer ML, Willson D, et al: Defining acute lung disease in children with the oxygen saturation index. *Pediatr Crit Care* Med 2010; 11:12–17
20. Ghuman AK, Newth CJ, Khemani RG: The association between the end tidal alveolar dead space fraction and mortality in pediatric acute hypoxemic respiratory failure. *Pediatr Crit Care Med* 2012; 13:11–15
21. Epstein D, Wong CF, Khemani RG, et al: Race/ethnicity is not associated with mortality in the PICU. *Pediatrics* 2011; 127:e588–e597
22. Bland JM, Altman DG, Bland JM, et al: Agreement between methods of measurement with multiple observations per individual. *J Biopharm Stat* 2007; 17:571–582
23. Hosmer D, Lemeshow S: Applied Logistic Regression. Second Edition. New York, Wiley-Interscience, 2000
24. 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
25. Khemani RG, Newth CJL: The design of future pediatric mechanical ventilation trials for acute lung injury. *Am J Respir Crit Care Med* 2010; 182:1465–1474