The S/F Illusion:

Can we define ARDS without blood gases?

Demetrio Magatti

Istituto di Ricerche Farmacologiche Mario Negri IRCCS

Tommaso Tonetti

Alma Mater Studiorum Università di Bologna
IRCCS Azienda Ospedaliero-Universitaria Bologna

The Global definition of ARDS (2023)

WORKSHOP

A New Global Definition of Acute Respiratory Distress Syndrome

Michael A. Matthay[1,2,3], Yaseen Arabi[5], Alejandro C. Arroliga[6], Gordon Bernard[7], Andrew D. Bersten[9], Laurent J. Brochard[10], Carolyn S. Calfee[1,2,3], Alain Combes[11], Brian M. Daniel[4], Niall D. Ferguson[12,13], Michelle N. Gong[14], Jeffrey E. Gotts[15], Margaret S. Herridge[16], John G. Laffey[17], Kathleen D. Liu[1,2], Flavia R. Machado[18], Thomas R. Martin[19], Danny F. McAuley[20], Alain Mercat[21], Marc Moss[22], Richard A. Mularski[23], Antonio Pesenti[24,25], Haibo Qiu[26], Nagarajan Ramakrishnan[27], V. Marco Ranieri[28], Elisabeth D. Riviello[29], Eileen Rubin[30], Arthur S. Slutsky[10], B. Taylor Thompson[31], Theogene Twagirumugabe[32], Lorraine B. Ware[8], and Katherine D. Wick[33]

Berlin Definition	Rationale for Updating Criteria	How This is Addressed in the Global Definition
Acute onset within 1 week of known insult or new or worsening respiratory symptoms	Onset may be more indolent for some insults, such as COVID-19	The inclusion of patients with HFNO will capture patients with more indolent courses, and therefore the timing criterion has not been changed
Bilateral opacities on chest radiography or computed tomography not fully explained by effusions, lobar/lung collapse, or nodules	Chest radiography and computed tomography not available in some clinical settings	Ultrasound can be used to identify bilateral loss of lung aeration (multiple B lines and/or consolidations) as long as operator is well trained in the use of ultrasound
Three severity categories defined by $Pa_{O_2}:FI_{O_2}$	Pulse oximetric measurement of $Sp_{O_2}:FI_{O_2}$ is widely used and validated as a surrogate for $Pa_{O_2}:FI_{O_2}$	$Sp_{O_2}:FI_{O_2}$ can be used for diagnosis and assessment of severity if Sp_{O_2} is ≤97%
Requirement for invasive or noninvasive mechanical ventilation such that PEEP ≥ 5 cm H_2O is required for all categories of oxygenation severity except mild, which can also be met with CPAP ≥ 5 cm H_2O	HFNO increasingly being used in patients with severe hypoxemia who otherwise meet ARDS criteria Invasive and noninvasive mechanical ventilation not available in resource-limited settings	New category of nonintubated ARDS created for patients on HFNO at ≥30 L/min who otherwise meet ARDS criteria Modified definition of ARDS for resource-limited settings does not require $Pa_{O_2}:FI_{O_2}$, PEEP, or HFNO

Should we define ARDS in «less intensive» settings?

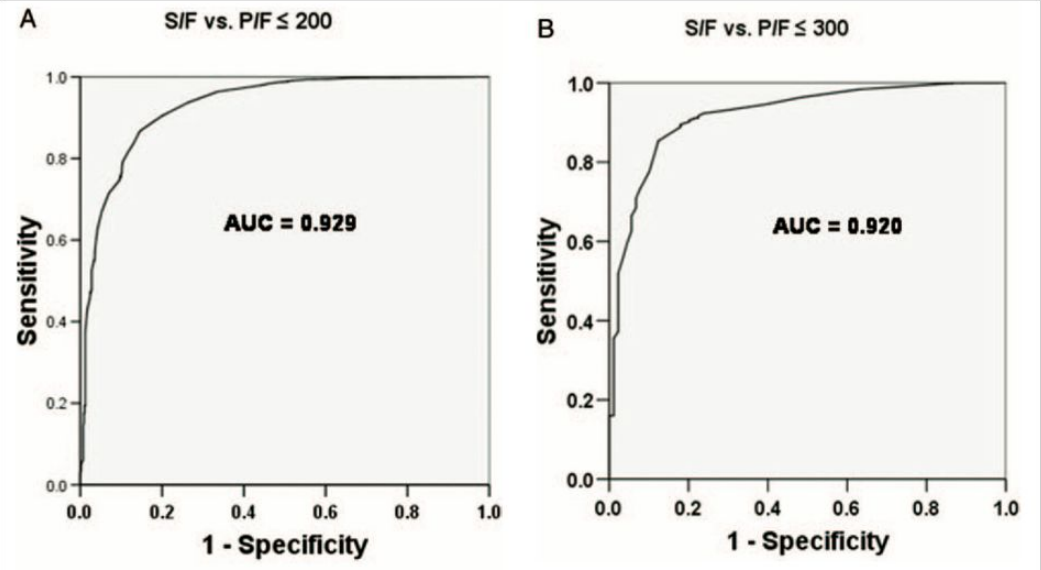
Possible reasons:

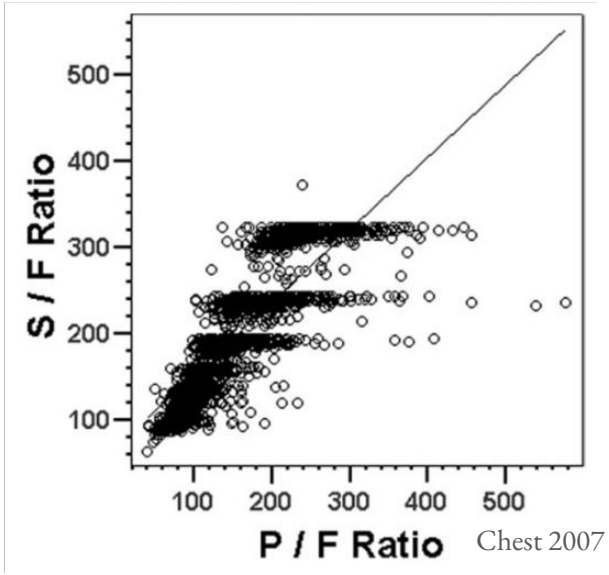
- For clinical reasons
  - Early identification of at-risk patients?
  - Early treatment in ICU?
- For research purpose
  - Epidemiology?
  - Early enrollment in clinical trials?
  - Studies on prevention of the syndrome?
- For lack of resources?

Is the S/F ratio a **good enough** tool for these purposes?

Comparison of the Spo_2/Fio_2 Ratio and the Pao_2/Fio_2 Ratio in Patients With Acute Lung Injury or ARDS*

Todd W. Rice, MD, MSc; Arthur P. Wheeler, MD, FCCP; Gordon R. Bernard, MD, FCCP; Douglas L. Hayden, MA; David A. Schoenfeld, PhD; and Lorraine B. Ware, MD, FCCP; for the National Institutes of Health, National Heart, Lung, and Blood Institute ARDS Network





Chest 2007

Patient records in M3 database

N Patients = 138387

(PaO2, SpO2, FiO2) triplets satisfying **time delay criterion**

N Patients = 69333
N triplets = 334488

Data cleaning

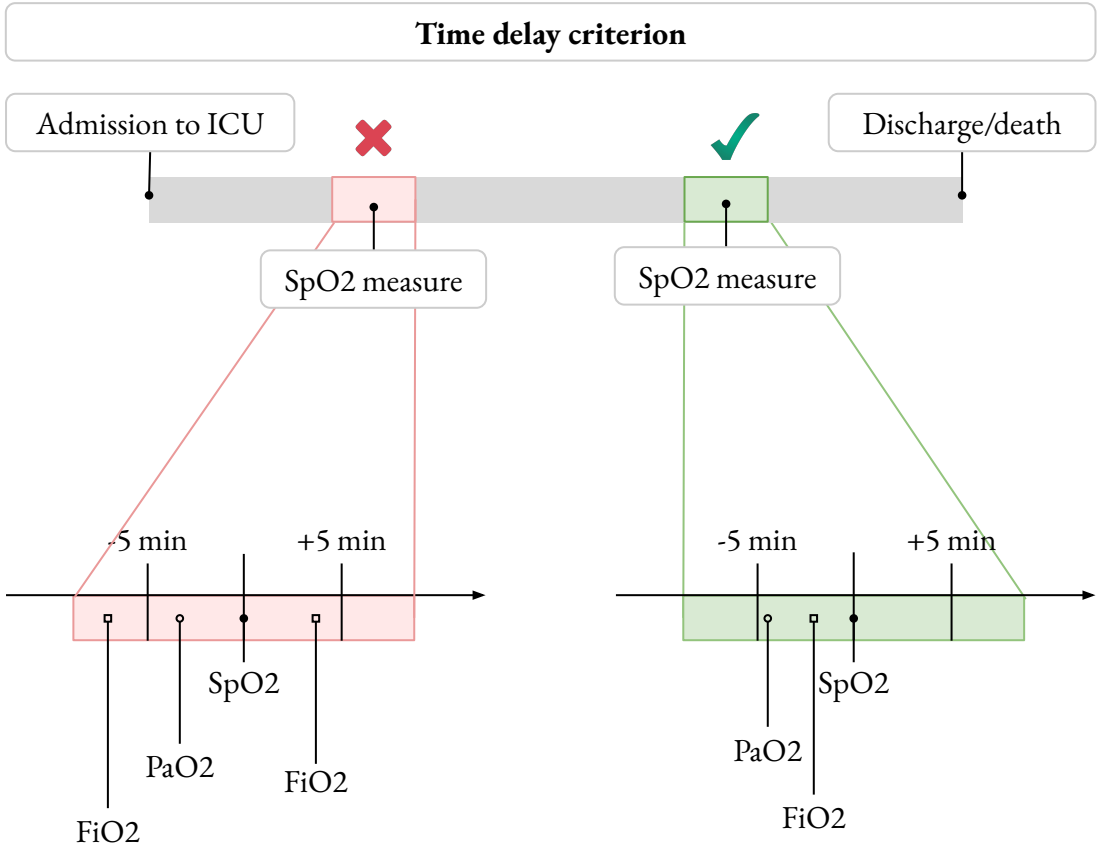
N triplets = 26483

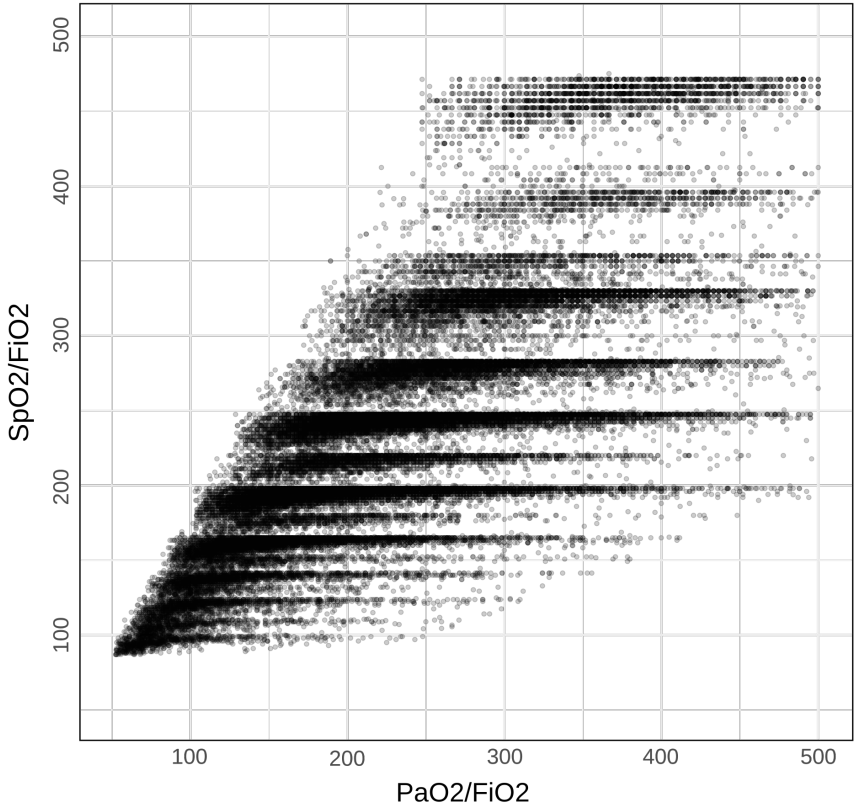
Clean data

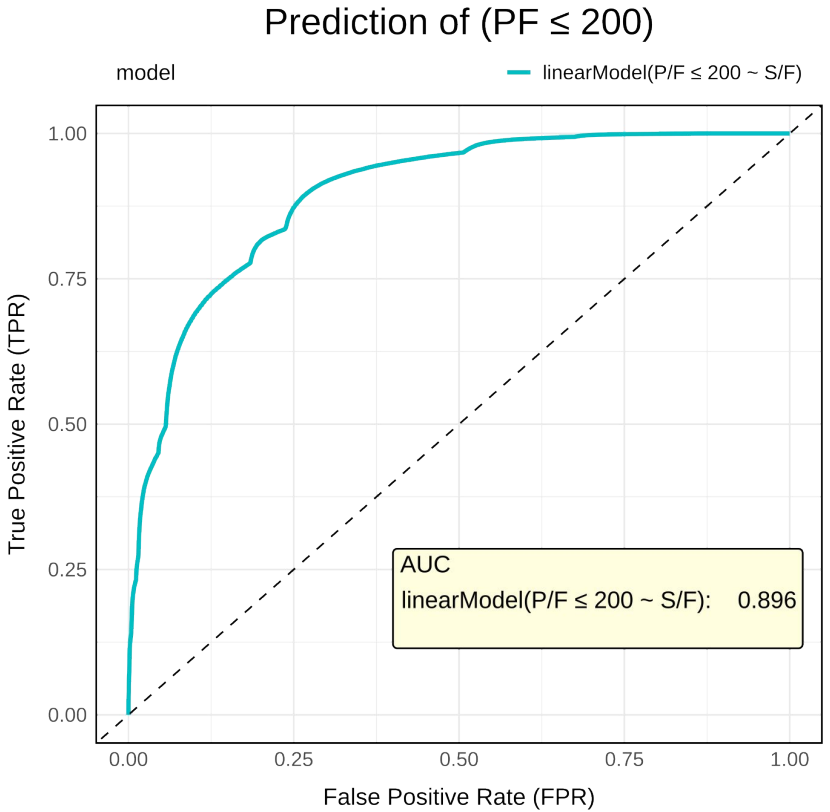
N Patients = 65315
N triplets = 308005

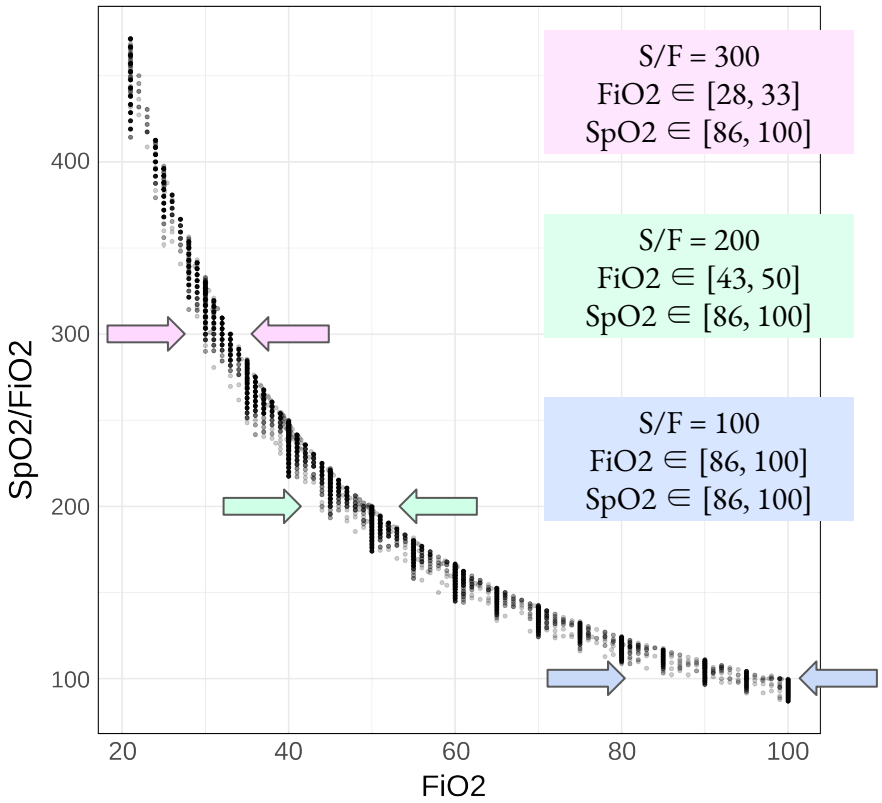
(PaO2, SpO2, FiO2) triplets with SpO2 ≤ 99%

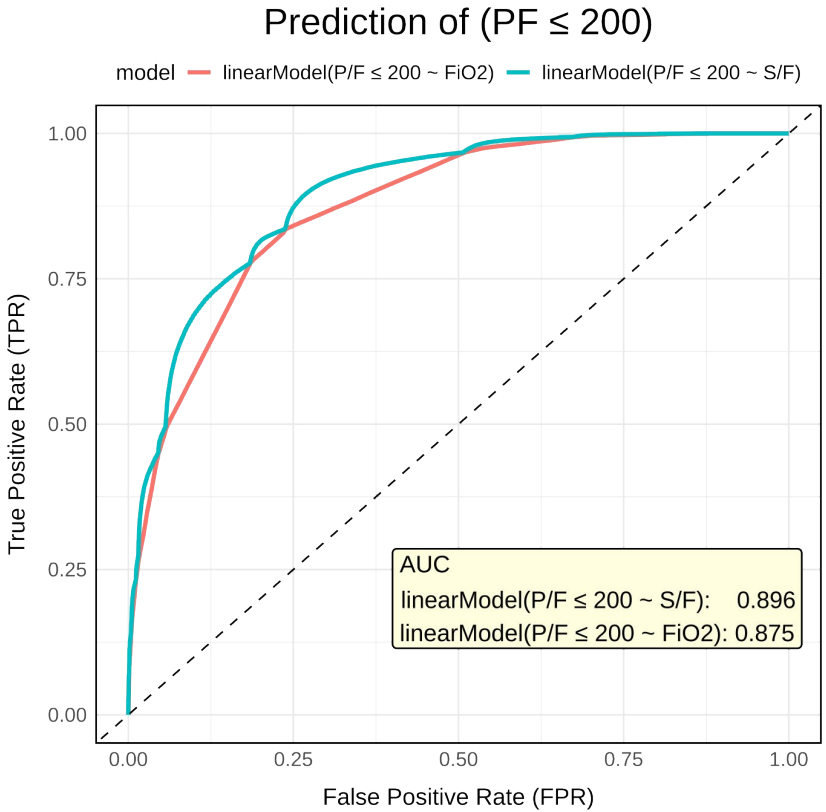
N Patients = 53562
N triplets = 209003

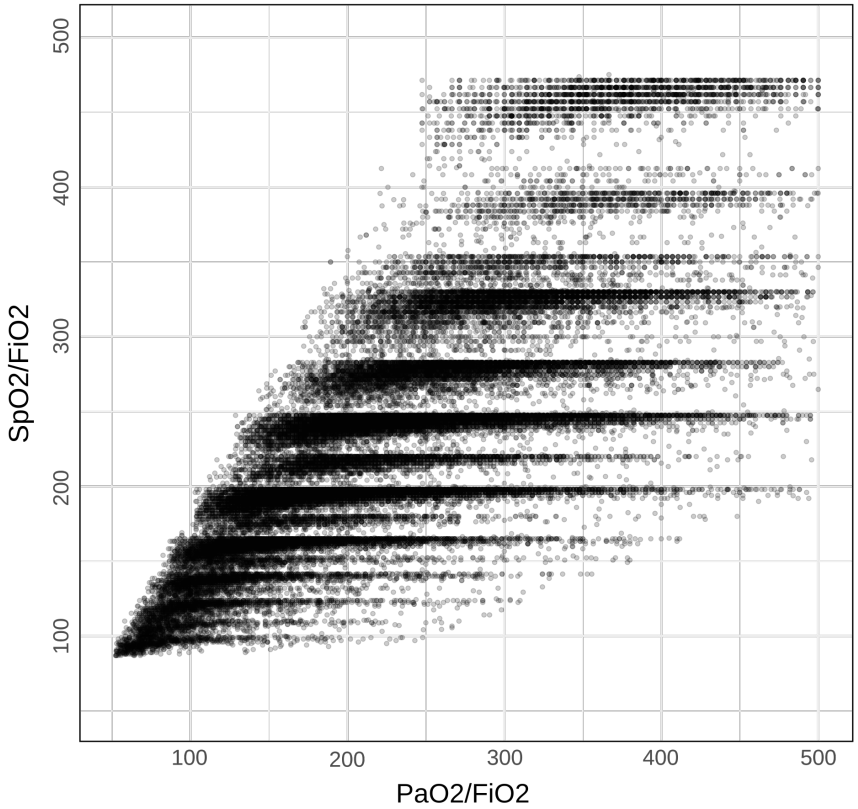


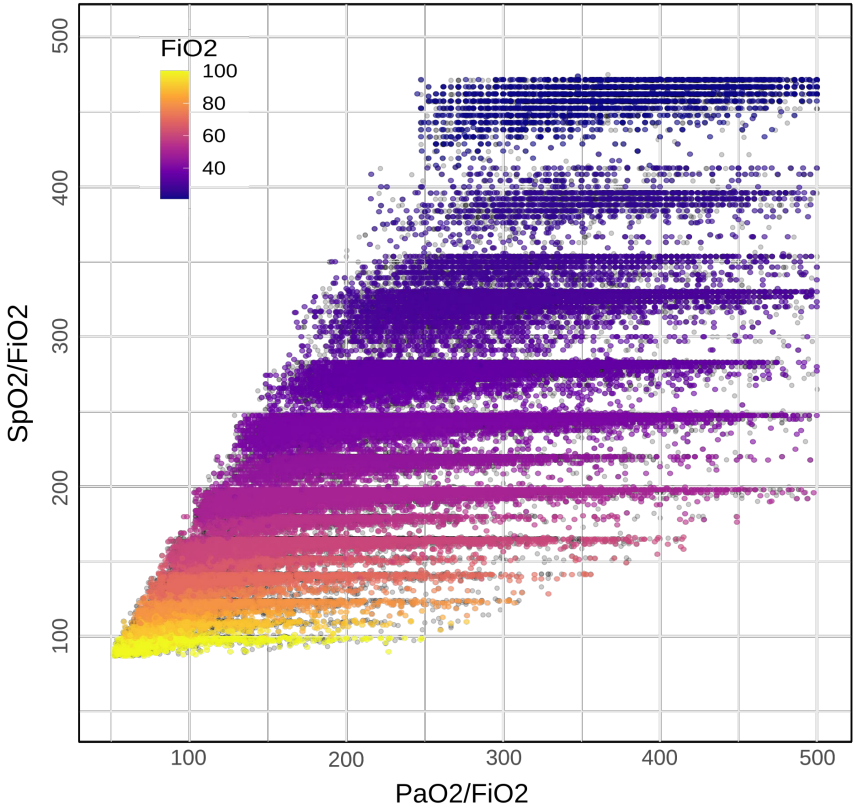

SpO2/FiO2
PaO2/FiO2
100
200
300
400
500

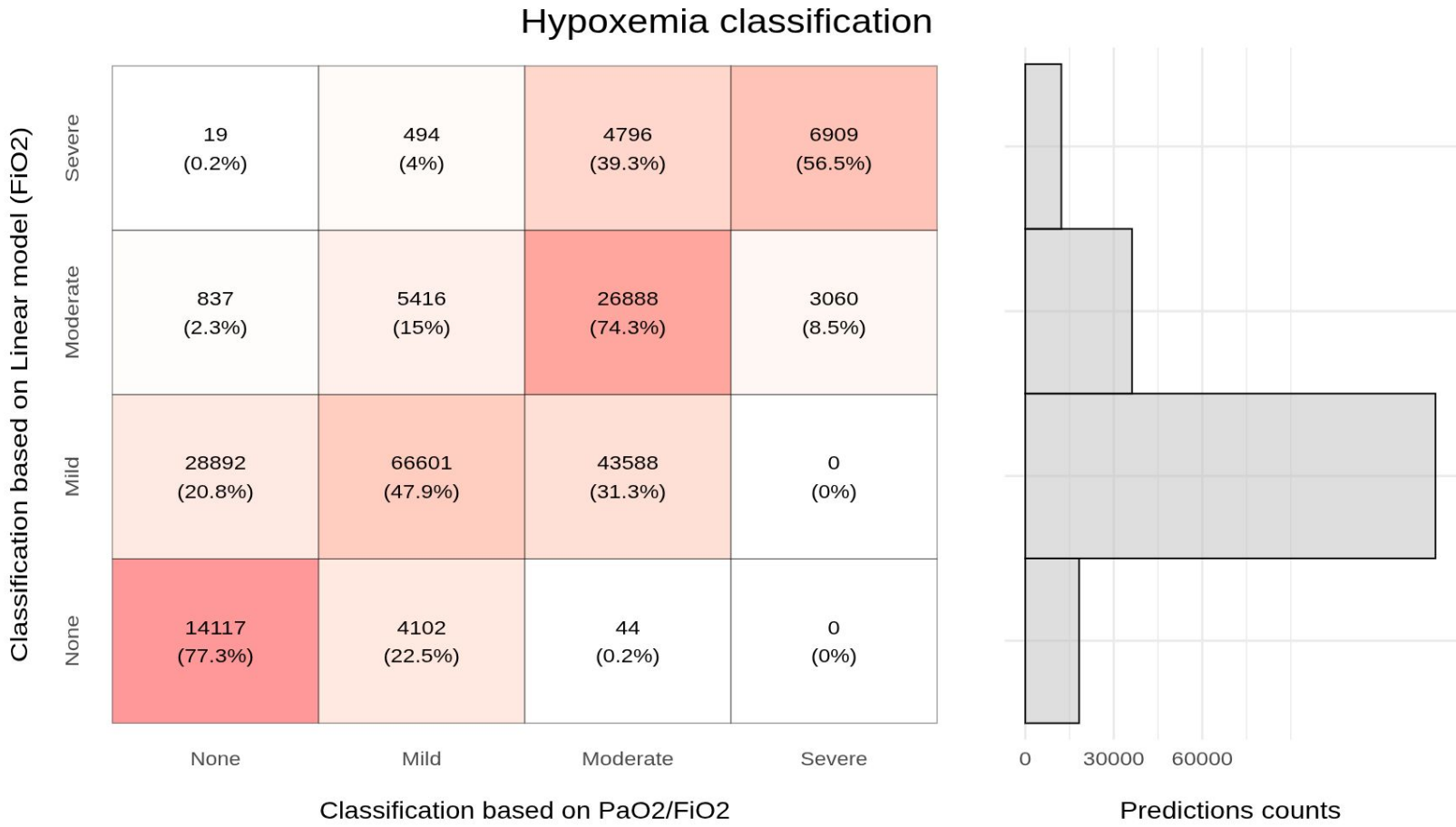

Prediction of (PF ≤ 200)
model
linearModel(P/F ≤ 200 ~ S/F)
True Positive Rate (TPR)
False Positive Rate (FPR)
0.00
0.25
0.50
0.75
1.00
AUC
linearModel(P/F ≤ 200 ~ S/F): 0.896


S/F = 300
FiO2 ∈ [28, 33]
SpO2 ∈ [86, 100]
S/F = 200
FiO2 ∈ [43, 50]
SpO2 ∈ [86, 100]
S/F = 100
FiO2 ∈ [86, 100]
SpO2 ∈ [86, 100]
SpO2/FiO2
FiO2
100
200
300
400
20
40
60
80
100


Prediction of (PF ≤ 200)
model
linearModel(P/F ≤ 200 ~ FiO2)
linearModel(P/F ≤ 200 ~ S/F)
True Positive Rate (TPR)
False Positive Rate (FPR)
0.00
0.25
0.50
0.75
1.00
AUC
linearModel(P/F ≤ 200 ~ S/F): 0.896
linearModel(P/F ≤ 200 ~ FiO2): 0.875


SpO2/FiO2
PaO2/FiO2
100
200
300
400
500


FiO2
100
80
60
40
SpO2/FiO2
PaO2/FiO2
100
200
300
400
500

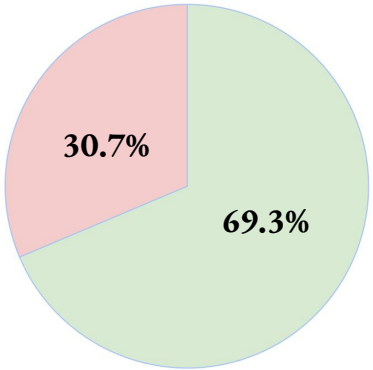

Hypoxemia classification
Classification based on Linear model (FiO2)
Severe
Moderate
Mild
None
19 (0.2%)
494 (4%)
4796 (39.3%)
6909 (56.5%)
837 (2.3%)
5416 (15%)
26888 (74.3%)
3060 (8.5%)
28892 (20.8%)
66601 (47.9%)
43588 (31.3%)
0 (0%)
14117 (77.3%)
4102 (22.5%)
44 (0.2%)
0 (0%)
None
Mild
Moderate
Severe
Classification based on PaO2/FiO2
0
30000
60000
Predictions counts

Results: linear model P/F ~ 1/FiO2				
	Class: None	Class: Mild	Class: Moderate	**Class: Severe**
Sensitivity	0,322	0,869	0,357	**0,693**
Precision	0,773	0,479	0,743	**0,565**

Sensitivity

Out of all *true* Moderate cases, how many did the model correctly predict as Moderate?

High sensitivity = fewer false negatives



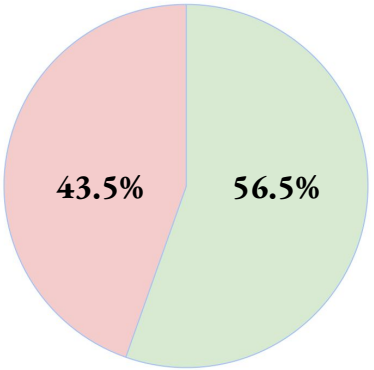
Of all cases that were actually "Severe", 69.3% have been correctly identified se "Severe" by the model

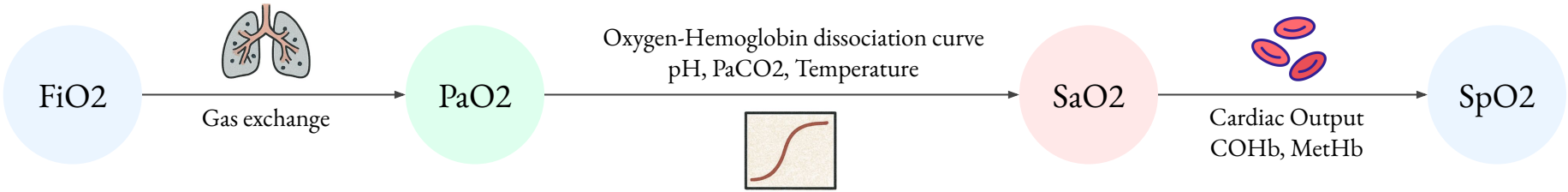
Precision

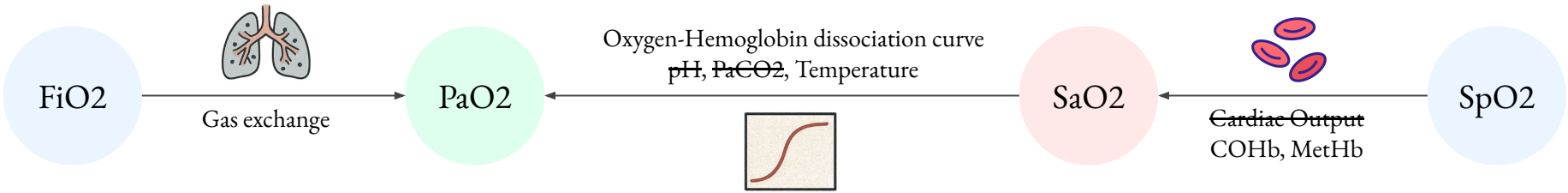
When the model predicts "Moderate," how often is it actually correct?

High precision= reliable positive predictions

Of all cases that were predicted "Severe", 56.5% were actually "Severe" according to classification based on P/F




FiO2
Gas exchange
PaO2
Oxygen-Hemoglobin dissociation curve
pH, PaCO2, Temperature
SaO2
Cardiac Output
COHb, MetHb
SpO2


FiO2
Gas exchange
PaO2
Oxygen-Hemoglobin dissociation curve
~~pH~~, ~~PaCO2~~, Temperature
SaO2
~~Cardiac Output~~
COHb, MetHb
SpO2

	Linear model P/F ~ S/F		Non linear model PaO2 ~ s(FiO2) + s(SpO2)		
	Sensitivity	Precision	Sensitivity	Precision	Prevalence
Class: None	0,535	0,698	0,534	0,720	0,213
Class: Mild	0,623	0,572	0,752	0,579	0,372
Class: Moderate	0,746	0,637	0,685	0,761	0,366
Class: Severe	0	NA	0,511	0,854	0,048

Sensitivity: out of all *true* Moderate cases, how many did the model correctly predict as Moderate?

High sensitivity = fewer false negatives

Precision: when the model predicts "Moderate," how often is it actually correct?

Precision = reliable positive predictions

	Linear model P/F ~ S/F		Non linear model PaO2 ~ s(FiO2) + s(SpO2)		
	Sensitivity	Precision	Sensitivity	Precision	Prevalence
Class: None	0,535	0,698	0,534	0,720	0,213
Class: Mild	0,623	0,572	0,752	0,579	0,372
Class: Moderate	0,746	0,637	0,685	0,761	0,366
Class: Severe	0	NA	0,511	0,854	0,048

Conclusions

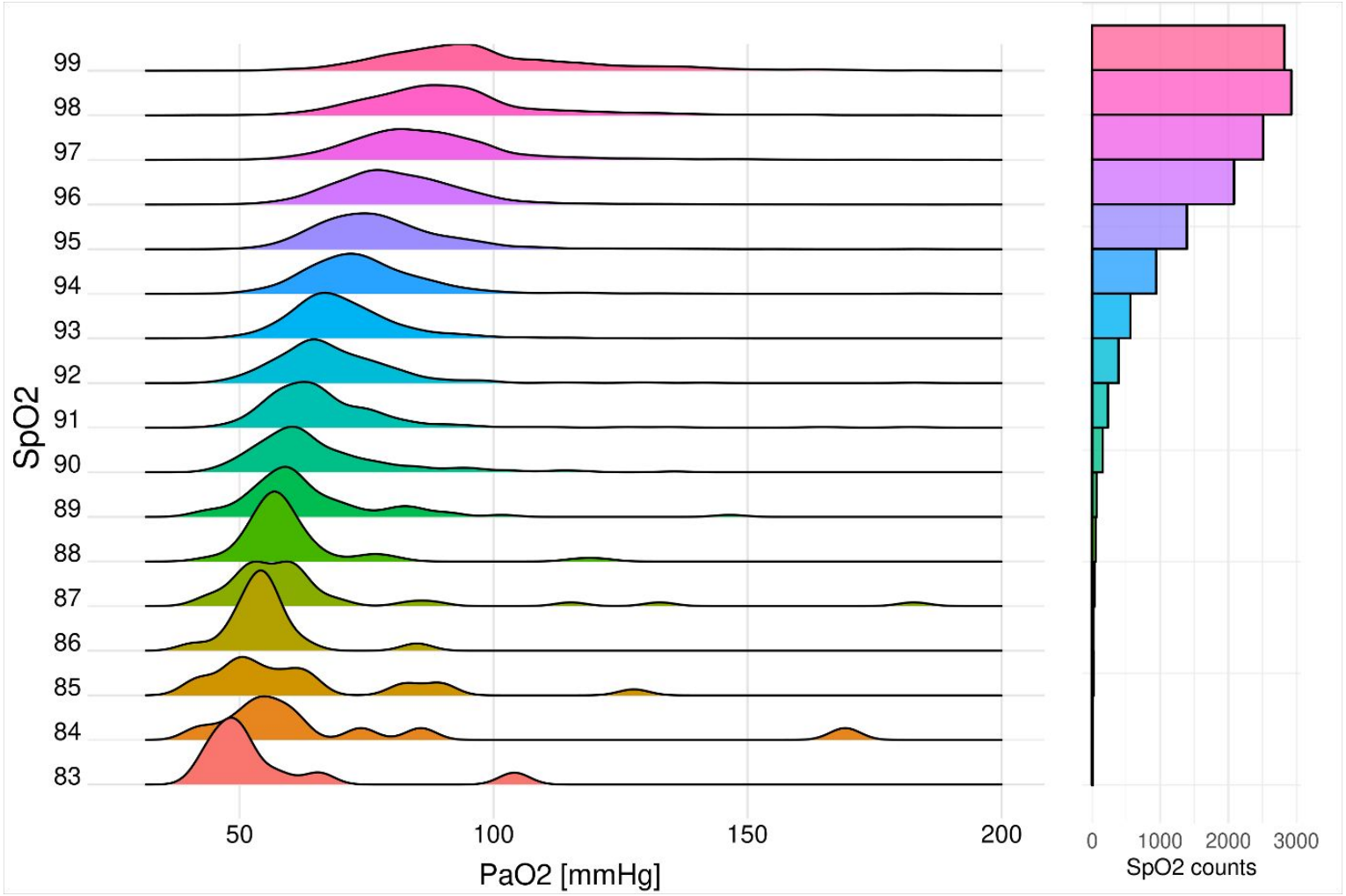
- S/F is a practical but imperfect surrogate of P/F.
- The relationship is nonlinear and affected by multiple physiological factors.
- Linear models show moderate accuracy; nonlinear ones perform better but remain limited.
- Useful for screening and triage, not for ARDS definition.
- Defining ARDS still means integrating physiology, imaging, and clinical context.

Perspectives

- Develop nonlinear, data-driven models (e.g., GAM, ML) including more variables.
- Validate prospectively in less intensive settings to detect early ARDS risk.
- Integrate into monitoring systems or EHRs for early alerts.
- Harmonize classification criteria between global definition and clinical practice.
- Move from static ratios to dynamic, physiology-based assessment of gas exchange.

Grazie per l’attenzione

Backup slides


SpO2
99
98
97
96
95
94
93
92
91
90
89
88
87
86
85
84
83
50
100
150
200
PaO2 [mmHg]
0
1000
2000
3000
SpO2 counts

Sensitivity: out of all *true* Moderate cases, how many did the model correctly predict as Moderate? High sensitivity = fewer false negatives

Specificity: out of all cases that are *not* Moderate, how many did the model correctly avoid labeling as Moderate? High specificity = fewer false positives

Pos Pred Value: when the model predicts "Moderate," how often is it actually correct? High PPV = reliable positive predictions

Neg Pred Value: when the model predicts something *other than* Moderate, how often is it correct?

Prevalence: proportion of specfic class in the dataset.

Results: linear model P/F ~ 1/FiO2

	Sensitivity	Specificity	Pos Pred Value	Neg Pred Value	Prevalence
Class: None	0,322	0,974	0,773	0,841	0,213
Class: Mild	0,869	0,438	0,479	0,850	0,372
Class: Moderate	0,357	0,929	0,743	0,714	0,366
Class: Severe	0,693	0,973	0,565	0,984	0,048

Sensitivity: out of all *true* Moderate cases, how many did the model correctly predict as Moderate? High sensitivity = fewer false negatives

Specificity: out of all cases that are *not* Moderate, how many did the model correctly avoid labeling as Moderate? High specificity = fewer false positives

Pos Pred Value: when the model predicts "Moderate," how often is it actually correct? High PPV = reliable positive predictions

Neg Pred Value: when the model predicts something *other than* Moderate, how often is it correct?

Prevalence: proportion of specfic class in the dataset.

Results: nonlinear model P/F ~ s(FiO2) + s(SpO2)

	Sensitivity	Specificity	Pos Pred Value	Neg Pred Value	Prevalence
Class: None	0,534	0,944	0,720	0,882	0,213
Class: Mild	0,752	0,676	0,579	0,821	0,372
Class: Moderate	0,685	0,876	0,761	0,828	0,366
Class: Severe	0,511	0,996	0,854	0,976	0,048