Title Page

Development and validation of a point-of-care assay to identify ARDS subphenotypes

Reddy K[1,2], Antcliffe D[3,4], Sinha P[5], Conlon J[1], McDowell C[6], Delucchi K[7], Calfee CS[8,9], O'Kane CM[1], McAuley DF[1,2]

1. Wellcome-Wolfson Institute for Experimental Medicine, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, 97 Lisburn Rd, Belfast BT9 7BL

2. Regional Intensive Care Unit, Royal Victoria Hospital, 274 Grosvenor Road, Belfast, BT12 6BA

3. Division of Anaesthetics, Pain Medicine & Intensive Care, Department of Surgery & Cancer, Faculty of Medicine, Imperial College London, London, UK

4. Department of Critical Care, Imperial College Healthcare NHS Trust, London, UK

5. Division of Critical Care, Department of Anesthesiology, Washington University in St Louis, St Louis, Missouri, United States

6. Clinical Trials Unit, Royal Victoria Hospital, 274 Grosvenor Road, Belfast, BT12 6BA

7. Department of Psychiatry, University of California, San Francisco, San Francisco, CA

8. Department of Medicine, Division of Pulmonary, Critical Care, Allergy and Sleep Medicine; University of California San Francisco, San Francisco, CA

9. Department of Anaesthesia, University of California San Francisco, San Francisco, CA

Word count: Abstract 344, Manuscript 2042

Tables: 2

Figures: 3

Abstract

Background

ARDS has no effective pharmacological therapy. Retrospective analyses of patients with ARDS in clinical trials have identified “hypoinflammatory” and “hyperinflammatory” subphenotypes based on a combination of clinical data and cytokine data measured by traditional ELISA. These subphenotypes had differential response to treatments in retrospective analyses. Inflammatory subphenotype stratification could therefore enable a precision medicine approach in ARDS, but requires prospective identification of the phenotype. We aimed to test whether a point-of-care (POC) assay could be used to measure the cytokines to allow real-time stratification to subphenotype.

Aims

To compare the classification of patients with ARDS as hypo- or hyperinflammatory using cytokine data derived from a POC assay with data from traditional ELISA.

Methods

A rapid assay for interleukin-6 (IL-6) and soluble tumour necrosis factor-1 (sTNFR1) was developed. This was used to measure IL-6 and sTNFR1 in 98 randomly selected plasma samples from ARDS patients enrolled in the HARP-2 clinical trial, and 165 patients with ARDS from the PHIND study. Hypo- or hyperinflammed subphenotype was allocated using a previously-described 3 variable parsimonious logistic regression model (IL-6, sTNFR1 and vasopressor use) using cytokine data from the POC assay and from laboratory gold standard ELISA. Allocation according to POC and ELISA data was compared.

Results

98 patients samples from HARP-2 and 165 samples from PHIND were analysed. Probability of subphenotype allocation for POC and ELISA data was highly concordant in HARP-2 ($r = 0.94$) and in PHIND ($r = 0.87$). In HARP-2, 92% of patients had concordant subphenotype allocation between POC and ELISA biomarker measurements, with no significant difference in disagreement (McNemar’s $\chi^2 = 0.13$; $p = 0.72$). In PHIND, 85% of patients had concordant subphenotype allocation between POC and ELISA biomarker

measurements with no significant difference in disagreement (McNemar's $\chi^2 = 0.36$; $p = 0.69$).

Conclusion

Allocation to hyper and hypoinflammatory subphenotypes using a parsimonious model including IL-6, sTNFR1 and vasopressor use is similar when plasma cytokines are measured using a novel POC assay compared with gold standard laboratory ELISA. These data support use of the POC assay for prospective phenotyping of patients with ARDS.

Background

10.4% of patients admitted to ICU meet criteria for the acute respiratory distress syndrome (ARDS), which has a hospital mortality rate of 34.9% to 46.1% (1). Numerous randomised controlled trials (RCTs) of pharmacological therapies in ARDS have failed to show treatment benefit (2). The only trials showing benefit for all-comers with ARDS have been for supportive therapies, such as careful ventilator and fluid management strategies (3-5). ARDS is a syndromic definition, encompassing multiple different aetiologies and associated with wide heterogeneity in underlying biological processes. This heterogeneity is believed to be responsible to the failure to find specific pharmacological therapies that are effective in all-comers. It is probable that, under the umbrella of the current Berlin definition of ARDS (6), both responders and non-responders are being recruited to clinical trials, resulting in a net result of no benefit for the overall cohort. Even more worryingly, a subgroup of patients may experience harm as a result of trial designs that recruit all-comers with ARDS.

In efforts to tease out the underlying heterogeneity and to allow focus on potential responders (predictive enrichment), two subphenotypes of ARDS have been identified through latent class analysis (LCA) of clinical and biomarker data from previous RCTs (7). The hyperinflammatory subphenotype was characterized by higher circulating levels of inflammatory markers (interleukin-6 [IL-6], interleukin-8 [IL-8], soluble tumour necrosis factor receptor-1 [sTNFR1], and plasminogen activator inactivator-1 [PAI-1]), lower Protein C, more use of vasopressors, more metabolic acidosis, and a greater prevalence of sepsis. The hypoinflammatory subphenotype was more prevalent (67-74% vs. 26-33%) and had lower mortality (19-23% vs. 44-51%). These results have been replicated in repeated retrospective analyses of clinical cohorts with a high degree of concordance, and the subphenotypes showed differential mortality in response to ventilation strategies (7), fluid-management strategies (8), and statins (9), thereby pointing towards underlying mechanistic differences. If subphenotypes can be allocated at the bedside, they could potentially be used to stratify patients for precision medicine RCTs, where treatments are targeted to subphenotypes that

are expected to be more treatment-responsive. However, an LCA approach to identifying the subphenotypes requires numerous laboratory biomarker measurements and population-level datasets, making it impractical for clinical use.

We have previously developed parsimonious logistic regression models that use a limited set of plasma cytokine measurements as well as clinical data on vasopressor use and /or plasma bicarbonate to identify the hypoinflammatory and hyperinflammatory subphenotypes with a high degree of accuracy compared with the gold standard LCA (10). Such parsimonious models have been derived and validated in data from four ARDS network RCTs, the HARP-2 trial of simvastatin in ARDS, and the START trial of mesenchymal stromal cells in ARDS (3, 5, 11-13). The logistic regression models use a probability cut-off of equal to or greater than 0.5 to allocate patients to the hyperinflammatory subphenotype, with those patients with lower probabilities being allocated to the hypoinflammatory subphenotype.

Prospective subphenotype determination is necessary to allow randomisation of patients to precision medicine trials. The parsimonious models could facilitate this, but this approach would require a rapid measurement of the relevant cytokines, which is not practical with traditional ELISAs. The aim of this work was to test whether a POC assay could be used to measure IL-6 and sTNFR1, and compare allocation to subphenotypes using data derived from the POC assay versus traditional ELISA in a parsimonious model.

Methods

Biological samples

HARP-2 was a randomised, controlled clinical trial evaluating simvastatin in 540 patients with ARDS (ISRCTN88244364). Patients were recruited within 48 hours of fulfilling criteria for ARDS and plasma collected in Lithium heparin tubes at time of recruitment (12). Plasma was separated by centrifugation of whole blood at the recruiting site before freezing for further analysis.

PHIND is a currently-recruiting multicentre observational study, aiming to recruit 480 ARDS patients (NCT04009330) within 72 hours of ARDS onset. Blood is collected in Lithium heparin tubes, before separation into plasma. Plasma is used to measure IL-6 and sTNFR1 via the POC assay at the bedside, and the remaining plasma sample frozen and stored at -80 ^{o}C for further analysis. As part of a planned interim analysis, a convenience sample of plasma samples from patients recruited to PHIND before July 2022 was used.

POC assay

A rapid quantitative plasma immunoassay with chemiluminescent signal detection for IL-6 and sTNFR1 was developed for use by Randox laboratories with their BioChip technology and Randox Evidence MultiSTAT (Randox Inc., Antrim, Northern Ireland) point-of-care quantitative analyser.

ELISA

IL-6 and sTNFR1 were measured on stored samples using DuoSet ELISA for IL-6 (DY206-05; R&D Systems, Bio-Techne, Minneapolis, USA) and Quantikine ELISA for sTNFR1 (DRT100; R&D Systems, Bio-Techne, Minneapolis, USA),

Statistical analysis

Several different parsimonious logistic regression models have been proposed with similar area under the receiver operating curves (AUCs) when compared with gold standard latent class analysis. For this work we chose a 3 variable parsimonious model encompassing IL-6, sTNFR1 and vasopressor use (since the HARP-2 study did not collect bicarbonate measurement). For both the HARP-2 and PHIND datasets, probability of allocation to the hyperinflammatory subphenotype was plotted for POC allocation versus ELISA allocation with the 3 variable parsimonious model and Pearson's correlation coefficient was calculated. A 2x2 table and McNemar's test were used to compare

subphenotype allocation between the POC assay and ELISA and this parsimonious model. Additionally, for the HARP-2 samples only, an area under the receiver operating curve (AUC) was calculated for POC and ELISA subphenotype allocation versus the original LCA classification. As LCA cannot be performed in the PHIND dataset until study completion, we were unable to generate AUC analysis in this cohort.

All analyses were performed using STATA software (version 15.1, StataCorp, College Station, Texas, USA).

Results

HARP-2 subphenotype allocation via parsimonious model using POC assay and ELISA data

98 patient samples were randomly selected from the HARP-2 study and run on the POC assay. Classification using the 3 variable parsimonious model was compared using ELISA and POC assay.

There was a strong linear correlation (r = 0.94) between the probability of hyperinflammatory subphenotype allocation using this parsimonious model with POC data versus ELISA data (Figure 1).

92% of patients (90/98) had matching subphenotype allocations as generated by this model for POC data versus ELISA data (Table 2). McNemar's chi-squared was 0.13 (p = 0.72), indicating the performance of the parsimonious model was equivalent when using POC data versus ELISA data.

PHIND subphenotype allocation via parsimonious model using POC assay and ELISA data

165 adult patients from PHIND with ARDS were included in this analysis. There was a strong linear correlation (r = 0.87) between probability of hyperinflammatory subphenotype allocation with this model using POC data versus ELISA data (Figure 2).

85% of patients (140/165) had matching subphenotype allocations as generated by this model for POC data versus ELISA data (Table 3). McNemar's chi-squared was 0.36 (p =

0.69), indicating the performance of this parsimonious model was equivalent when using POC data versus ELISA data.

HARP-2 subphenotype allocation via parsimonious model as compared to LCA

Allocation of patients in HARP-2 to the hyperinflammatory and hypoinflammatory subphenotypes was performed in previous work using gold-standard LCA, based on cytokine measurements using ELISAs (14). Data for similar analysis in PHIND is unfortunately not available yet, as LCA cannot be performed in this cohort until study completion.

In HARP-2, subphenotype allocation with ELISA data + this parsimonious model was strong, with AUC 0.94, confirming previous data that this parsimonious model can usefully phenotype patients (Figure 3). Subphenotype allocation with the POC assay was similar with AUC 0.89 using LCA subphenotype allocation as the gold-standard.

Discussion

This is the first demonstration that a point of care assay can be used to measure IL-6 and sTNFR1 rapidly, allowing potential prospective identification of inflammatory subphenotypes in ARDS alongside routinely available clinical data. The ability to use data from the POC assay in the parsimonious model to provide similar class allocation as using traditional ELISA data was shown by a post-hoc analysis of an RCT (HARP-2) and through interim analysis of an ongoing multicentre observational study (PHIND). We were able to demonstrate strong AUCs for both the ELISA and POC assay data in the parsimonious model against the gold standard LCA for HARP-2, and show high concordance in the allocations via the parsimonious model using both sets of cytokines measurements.

These data demonstrate that the subphenotypes can be accurately identified both retrospectively and prospectively using the POC assay. They also demonstrate that the subphenotypes exist both within the highly-selected patient cohort of a clinical trial, as well as within a more heterogeneous population such as that in PHIND. The POC assay takes

only 36 minutes to run, was deliverable by research nursing staff in the ICU, and produces subphenotype allocations that are highly consistent with ELISA and LCA data, supporting the concept that it can be used to stratify patients for randomisation to a precision medicine trial.

It is worth noting that a proportion of patients had discordant subphenotype allocations from POC and ELISA data. Some discrepancy is to be expected, as a perfect diagnostic test does not exist. In the original data in which the parsimonious models were derived, this parsimonious model had a sensitivity of 0.89 and sensitivity of 0.74 at a probability cutoff of 0.5 (10). In PHIND, since a complete LCA will not be undertaken until the end of the study, we have no gold-standard against which to compare the subphenotype allocations generated and displayed in Figure 2. However, the number of patients here with discordant allocations suggests that the performance in the PHIND dataset is likely not dissimilar to the stated sensitivities and sensitivities. Furthermore, it is reassuring that most of the patients with discordant subphenotype allocations lie somewhere in the middle of the probability distributions for both POC and ELISA measurements. These are the patients who are on the borderline of hypoinflammatory versus hyperinflammatory subphenotype allocation, and hence lie on the steepest part of the logistic regression curve. This minority of patients are relatively indeterminate by both POC and ELISA methods, suggesting that this is not measurement error, and in fact that this is a group that is difficult to stratify.

This raises a question as to potential overlap of the hypoinflammatory and hyperinflammatory subphenotypes. Clearly, based on previous work, these subphenotypes are highly reproducible and respond differently to treatment in retrospective analyses (7, 14, 15). In these datasets, the majority of patients classify strongly as hypoinflammatory or hyperinflammatory. However, in all cases there are a small minority of patients in the middle “indeterminate zone”, suggesting that there may be a temporal or physiological continuum between the two subphenotypes that we do not yet fully understand. Perhaps in future precision medicine trials that utilise the POC assay, this group could be targeted as a third subphenotype. Alternatively, probability cut-off thresholds for suphenotype allocation could

be varied based on prior expectation of risk versus benefit in the indeterminate group. For example, if we are testing a treatment that is likely to provide benefit in the hyperinflammatory subphenotype but has a high risk of side effects, we could randomise only the most hyperinflammatory patients to this treatment based on choosing a higher probability cutoff (e.g. 0.8) to minimise the risk of side effects to non-responders. Conversely, a lower risk treatment could have a more permissive probability cut-off.

Despite the remaining questions about these subphenotypes, our data suggest that the POC assay is ready to be used in a precision medicine trial. A rational first step in this approach would be to first use it to stratify treatments that have previously shown marginal or no benefit in the overall ARDS cohort, but have strong biological rationale for benefit in a particular subgroup, such as simvastatin in the hyperinflammatory subphenotype (14), corticosteroids, or other immune modulating drugs such as tocilizumab and anti-TNF drugs. The primary advantage of such an approach would be that regulatory approvals could be obtained rapidly for drugs such as these that have already proven safe in the ARDS population.

Considerable interest in reclassifying intensive care syndromes has been generated recently, and we are on the cusp of a paradigm shift in the field towards a more biology-driven understanding of critical illness (16). This work is the first of many to come that will represent a significant advancement towards the goal of precision critical care medicine.

References

1. Bellani G, Laffey JG, Pham T, Fan E, Brochard L, Esteban A, et al. Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. JAMA. 2016;315(8):788-800. https://doi.org/10.1001/jama.2016.0291.
2. Saha R, Assouline B, Mason G, Douiri A, Summers C, Shankar-Hari M. Impact of differences in acute respiratory distress syndrome randomised controlled trial inclusion and exclusion criteria: systematic review and meta-analysis. Br J Anaesth. 2021;127(1):85-101. https://doi.org/10.1016/j.bja.2021.02.027.
3. Acute Respiratory Distress Syndrome Network, Brower RG, Matthay MA, Morris A, Schoenfeld D, Thompson BT, et al. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. N Engl J Med. 2000;342(18):1301-8. https://doi.org/10.1056/nejm200005043421801.
4. Acute Respiratory Distress Syndrome Network. Higher versus lower positive end-expiratory pressures in patients with the acute respiratory distress syndrome. N Engl J Med. 2004;351(4):327-36. https://doi.org/10.1056/NEJMoa032193.
5. Wiedemann HP, Wheeler AP, Bernard GR, Thompson BT, Hayden D, deBoisblanc B, et al. Comparison of two fluid-management strategies in acute lung injury. N Engl J Med. 2006;354(24):2564-75. https://doi.org/10.1056/NEJMoa062200.
6. ARDS Definition Task Force, Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, et al. Acute respiratory distress syndrome: the Berlin definition. JAMA. 2012;307(23):2526-33. https://doi.org/10.1001/jama.2012.5669.
7. Calfee CS, Delucchi K, Parsons PE, Thompson BT, Ware LB, Matthay MA. Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. Lancet Respir Med. 2014;2(8):611-20. https://doi.org/10.1016/s2213-2600(14)70097-9.
8. Famous KR, Delucchi K, Ware LB, Kangelaris KN, Liu KD, Thompson BT, et al. Acute respiratory distress syndrome subphenotypes respond differently to randomized fluid management strategy. Am J Respir Crit Care Med. 2017;195(3):331-8. https://doi.org/10.1164/rccm.201603-0645OC.
9. Calfee CS, Delucchi KL, Sinha P, Matthay MA, Hackett J, Shankar-Hari M, et al. Acute respiratory distress syndrome subphenotypes and differential response to simvastatin: secondary analysis of a randomised controlled trial. Lancet Respir Med. 2018;6(9):691-8. https://doi.org/10.1016/s2213-2600(18)30177-2.
10. Sinha P, Delucchi KL, McAuley DF, O'Kane CM, Matthay MA, Calfee CS. Development and validation of parsimonious algorithms to classify acute respiratory distress syndrome phenotypes: a secondary analysis of randomised controlled trials. Lancet Respir Med. 2020;8(3):247-57. https://doi.org/10.1016/s2213-2600(19)30369-8.
11. Brower RG, Lanken PN, MacIntyre N, Matthay MA, Morris A, Ancukiewicz M, et al. Higher versus lower positive end-expiratory pressures in patients with the acute respiratory distress syndrome. N Engl J Med. 2004;351(4):327-36. https://doi.org/10.1056/NEJMoa032193.
12. McAuley DF, Laffey JG, O'Kane CM, Perkins GD, Mullan B, Trinder TJ, et al. Simvastatin in the acute respiratory distress syndrome. N Engl J Med. 2014;371(18):1695-703. https://doi.org/10.1056/NEJMoa1403285.
13. Matthay MA, Calfee CS, Zhuo H, Thompson BT, Wilson JG, Levitt JE, et al. Treatment with allogeneic mesenchymal stromal cells for moderate to severe acute respiratory distress syndrome (START study): a randomised phase 2a safety trial. Lancet Respir Med. 2019;7(2):154-62. https://doi.org/10.1016/s2213-2600(18)30418-1.
14. Calfee CS, Delucchi KL, Sinha P, Matthay MA, Hackett J, Shankar-Hari M, et al. Acute respiratory distress syndrome subphenotypes and differential response to simvastatin: secondary analysis of a randomised controlled trial. Lancet Respir Med. 2018;6(9):691-8. https://doi.org/10.1016/S2213-2600(18)30177-2.

15. Famous KR, Delucchi K, Ware LB, Kangelaris KN, Liu KD, Thompson BT, et al. Acute respiratory distress syndrome subphenotypes respond differently to randomized fluid management strategy. Am J Respir Crit Care Med. 2017;195(3):331-8. https://doi.org/10.1164/rccm.201603-0645OC.
16. Maslove DM, Tang B, Shankar-Hari M, Lawler PR, Angus DC, Baillie JK, et al. Redefining critical illness. Nat Med. 2022;28(6):1141-8. https://doi.org/10.1038/s41591-022-01843-x.

Tables and Figures

Table 1: Concordance of subphenotype allocation in HARP-2 using the 3 variable parsimonious model (IL-6, sTNFR1, vasopressors) with POC cytokine versus ELISA cytokine measurements.

	ELISA hypoinflammatory	**ELISA hyperinflammatory**	**Total**
POC hypoinflammatory	63 (64%)	3 (3%)	66 (67%)
POC hyperinflammatory	5 (5%)	27 (28%)	32 (33%)
Total	68 (69%)	30 (31%)	98

Table 2: Concordance of subphenotype allocation in PHIND using the 3 variable parsimonious model (IL-6, sTNFR1, vasopressors) with POC cytokine versus ELISA cytokine measurements.

	ELISA hypoinflammatory	**ELISA hyperinflammatory**	**Total**
POC hypoinflammatory	102 (62%)	11 (7%)	113 (68%)
POC hyperinflammatory	14 (8%)	38 (23%)	52 (32%)
Total	116 (70%)	49 (30%)	165

Figure 1: Probability of allocation to the hyperinflammatory subphenotype in the HARP-2 dataset with the 3 variable parsimonious model (IL-6, sTNFR1, vasopressors) using point-of-care (POC) versus ELISA cytokine measurements. There is a strong linear correlation (r = 0.94). Hyperinflammatory subphenotype is allocated if probability ≥ 0.5 and hypoinflammatory subphenotype is allocated if probability < 0.5.

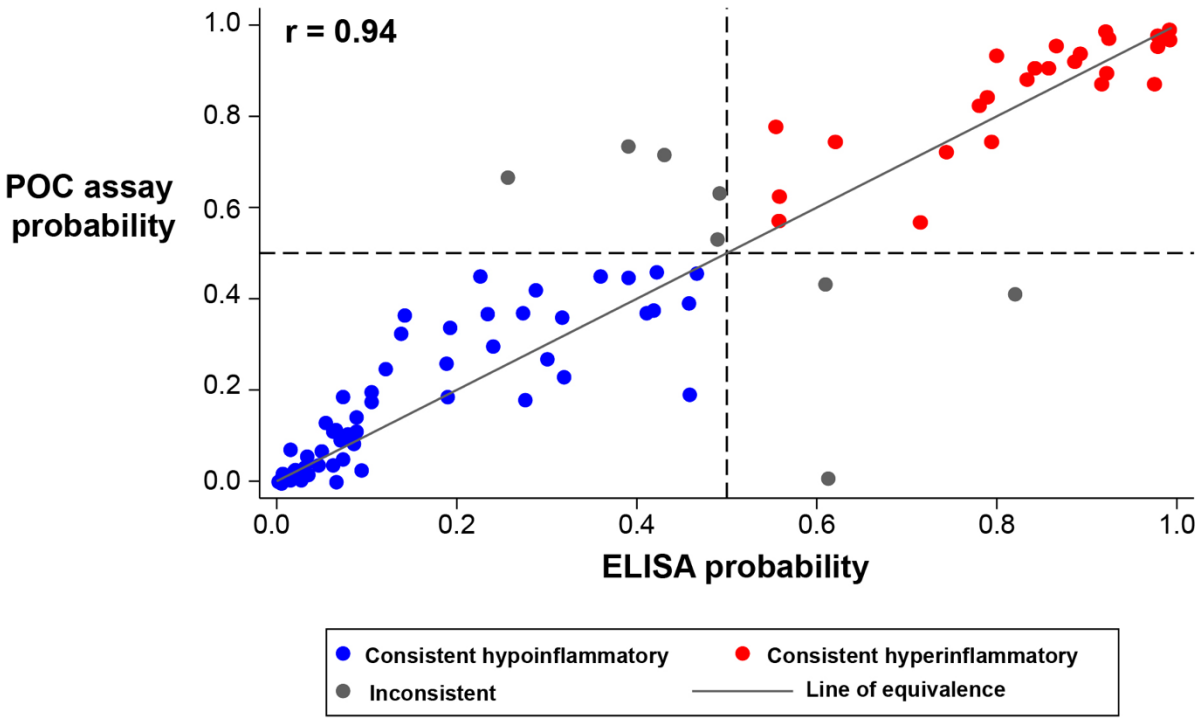

Figure 2: Probability of allocation to the hyperinflammatory subphenotype in the PHIND dataset with the 3 variable parsimonious model (IL-6, sTNFR1, vasopressors) using point-of-care (POC) versus ELISA cytokine measurements There is a strong linear correlation (r = 0.87). Hyperinflammatory subphenotype is allocated if probability ≥ 0.5 and hypoinflammatory subphenotype is allocated if probability < 0.5.

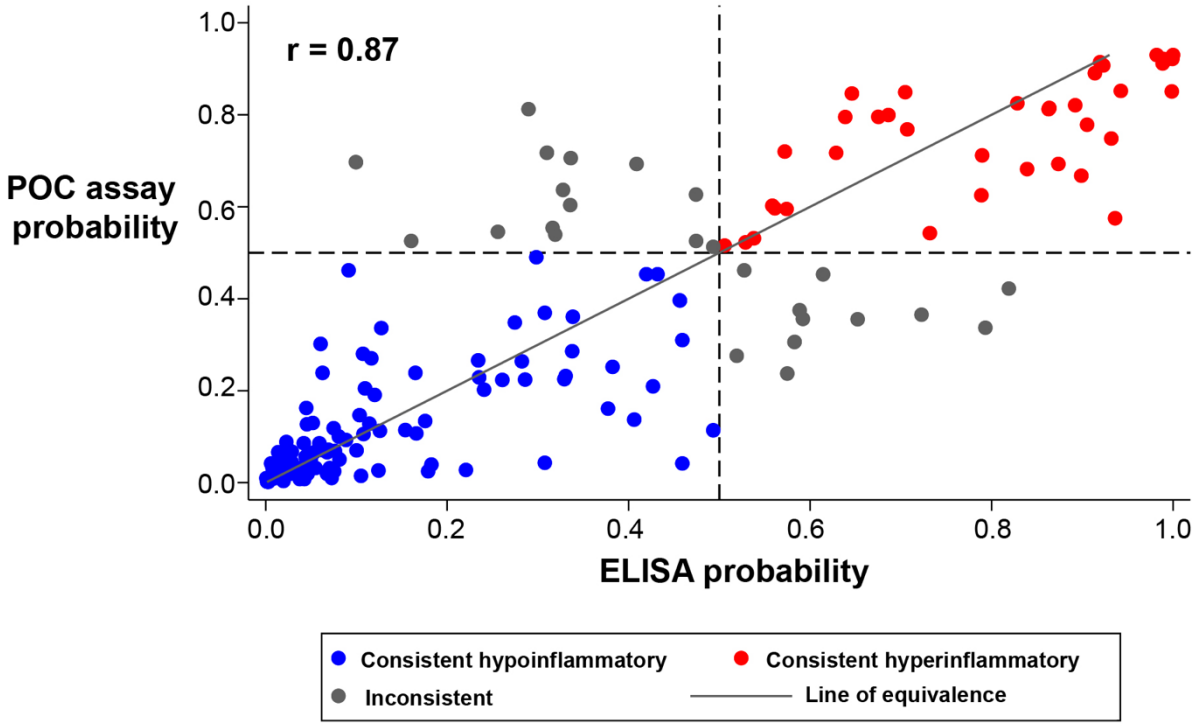

Figure 3: Area under the receiver operating curves (AUC) in the HARP-2 dataset using the 3 variable parsimonious model (IL-6, sTNFR1, vasopressors) and POC or ELISA cytokine measurements. Gold-standard was subphenotype allocation by latent class analysis (LCA).