*The Journal of Infectious Diseases*

**PROMISE SUPPLEMENT**

# External validation of the discriminative validity of the ReSVinet score & development of simplified ReSVinet scores in secondary care

**Zakariya Sheikh[1], Ellie Potter[1], You Li[2], Simon B Drysdale[3], Joanne G Wildenbeest[4], Hannah Robinson[5], Joseph McGinley[5], Gu-Lung Lin[5], Deniz Öner[6], Jeroen Aerssens[6], Antonio José Justicia-Grande[7], Federico Martinón-Torres[8], Andrew J Pollard[5], Louis Bont[4], Harish Nair[9]* on behalf of PROMISE investigators**

[1]Edinburgh Medical School, College of Medicine and Veterinary Medicine, University of Edinburgh, Edinburgh, UK.; [2]School of Public Health, Nanjing Medical University, Nanjing, China; [3]Centre for Neonatal & Paediatric Infection, St George's, University of London, London, UK; [4]Department of Pediatric Infectious Diseases and Immunology, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, The Netherlands.; [5]Oxford Vaccine Group, Department of Paediatrics, University of Oxford, and the NIHR Oxford Biomedical Research Centre, Oxford, United Kingdom; [6]Infectious Diseases Translational Biomarkers, Janssen Pharmaceutica NV, Beerse, Belgium; [7]Genetics, Vaccines and Infections Research Group (GENVIP). Instituto de Investigación Sanitaria de Santiago, University of Santiago, Santiago de Compostela, Spain.; [8]Department of Pediatrics, Translational Pediatrics and Infectious Diseases, Pediatrics Department, Hospital Clínico Universitario de Santiago de Compostela, Santiago de Compostela, Spain; [9]Usher Institute, College of Medicine and Veterinary Medicine, University of Edinburgh, Edinburgh, UK.

***Background:*** There is no consensus on how to best quantify disease severity in infants with respiratory syncytial virus (RSV) and/or bronchiolitis; this lack of a sufficiently validated score complicates the provision of clinical care and, the evaluation of trials of therapeutics and vaccines. The ReSVinet score appears to be one of the most promising; however it is too time-

*Corresponding author; Harish.Nair@ed.ac.uk



DOI: 10.1093/infdis/jiad388

consuming to be incorporated into routine clinical care. We aimed to develop and externally validate simplified versions of this score.

***Methods:*** Data were used from a multinational (Netherlands, Spain & United Kingdom) multicentre case-control observational study of infants with RSV to develop simplified versions of the ReSVinet by conducting a grid search to determine the best combination of equally weighted parameters to maximise for the discriminative ability of the scores across a range of outcomes (hospitalisation, intensive care unit admission, ventilation requirement). Subsequently discriminative validity of the score for a range of secondary care outcomes was externally validated by conducting a secondary analysis of data collected in infants with respiratory infection from tertiary hospitals in Rwanda and Colombia.

***Results:*** Three candidate simplified scores were identified using the development dataset; they were excellent (area under the receiver-operator characteristic curve [AUROC] >0.9) in the development dataset at discriminating for a range of outcomes, and their performance was not statistically significantly different to the original ReSVinet score despite having fewer parameters. In the external validation datasets, the simplified scores were moderate-excellent (AUROC 0.7-1) across a range of outcomes. In all outcomes, except for in a single dataset at predicting admission to the high dependency unit, they performed at least as well as the original ReSVinet score.

***Conclusions:*** Three promising candidate simplified scores were developed; however further external validation work in larger datasets, ideally from resource-limited settings needs to be conducted before any recommendation regarding their use.

Keywords: RSV, severity score, validity

## INTRODUCTION

There remains no consensus on how best to quantify the severity of bronchiolitis and/or respiratory syncytial virus (RSV) associated acute lower respiratory tract infections in infants; this is an important challenge as such a score can be useful indicator for clinicians when providing care and, serve as clinical endpoints in clinical trials for therapeutics and, increasingly importantly, vaccines. Of the various scores proposed, one of the most promising is the ReSVinet score [1-3].

The ReSVinet score was proposed almost a decade ago to quantify disease severity in infants (<24 months) with acute respiratory infections, including RSV [1]. It was designed based on an unreported literature review of previous related scores and subsequently evaluated by 90 paediatricians ensuring high face validity. The devised score consists of 7-parameters (feeding intolerance, medical intervention, respiratory difficulty, respiratory frequency, apnoea, general condition, fever) and has an overall total score of 20. The original developers of the score later

proposed thresholds (based on the findings from two studies) to indicate grades of severity: 0-6 as signifying a mild infection, 7-13 for moderate distress, and ≥ 14 for a severe episode [4]. Notably, this score doesn't include either oxygen saturation or heart rate, objective indicators often included in other RSV/bronchiolitis severity scores, making it more widely usable, especially in settings with limited equipment (e.g. outpatient and inpatient settings in low-income countries) [2]. A parental version of this score has also been developed, although the focus of this paper will be on the version for health professionals.

The original study that proposed the score retrospectively validated it in a small sample of 170 infants (<24 months) hospitalised in 3 centres in Spain with an acute respiratory infection; in this study they demonstrated that the score had good internal consistency, strong inter-rater reliability (between investigators and, between investigators and parents) and moderate construct validity [1]. Subsequently, a number of prospective external validation studies have been conducted showing similar results; studies assessing the validity of the clinician version of the ReSVinet score, that we are aware of, are summarised in *Supplementary Table 1* [1, 5-8]. As such it currently appears to be one of the most promising and best validated severity scores for acute respiratory infections in infants.

Anecdotally, the major reported weakness of the score is that it is too time-consuming to use, making it unfeasible for integrating it into routine clinical care (although still suitable for clinical trials). Therefore, we aimed to propose and externally validate a simplified version of the ReSVinet score. We additionally externally validated the original ReSVinet score in two new dataset.

## Methods

The design and reporting of this study was guided, as far as possible within the pre-existing constraints of the datasets, by the Prediction model Risk Of Bias ASsessment Tool (PROBAST) and the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) checklist [9,10]; for the detailed populated checklists see *Online supplementary file, appendix 1&2*.

### Data sources

For this study, we made use of three existing datasets. One recently published dataset was used to develop candidate simplified scores [11,12], and these scores were subsequently externally validated in three additional datasets, all of which have already been used to validate the original ReSVinet score and their findings published [5, 13]. Further detail on each dataset can be found below; the inclusion and exclusion criteria of the datasets are listed in *Supplementary Table 2*.

***Development dataset (RESCEU)***

The dataset used to develop simplified ReSVinet scores was collected as part of a recently published multinational (Netherlands, Spain & United Kingdom) multicentre case-control observational study run by the REspiratory Syncytial virus Consortium in EUrope (RESCEU) to investigate biomarkers of RSV associated acute lower respiratory tract infection severity in infants (<12 months) (NCT03756766) [12]. The original study aimed to recruit 500 previously healthy infants with RSV, 50 infants with comorbidities and with RSV and 80 healthy controls without RSV. Patients were recruited between October 2017 to April 2021.

For the purpose of this study, we only included the cases (i.e. infants with RSV), including those with comorbidities. The eligibility criteria for this group can be found in *Supplementary Table 2*. An anonymised version of this dataset was provided by the RESCEU investigators. For each patient, basic demographic data (e.g. sex, gestational age), clinical data (e.g. heart rate, oxygen saturation) outcome data (e.g. hospitalisation, need for respiratory support) and the ReSVinet score parameters and total score at the time of recruitment were supplied. Score assessors were not blinded for outcomes if the outcomes had occurred by the time of score assessment. Outcome data were retrieved from patients' medical notes; outcomes used in our analysis were hospitalisation, admission to the intensive care unit (ICU) and requirement for synchronised intermittent mandatory ventilation (SIMV).

***External validation datasets (Camacho-Cruz and Hakizimana)***

Subsequently we externally validated the simplified scores in two datasets, henceforth referred to as the Camacho-Cruz and Hakizimana datasets respectively. In this section, we provide a high-level summary of each dataset; however more detailed information on the methods can be found in the original study publications [5, 13].

The Camacho-Cruz dataset was collected in a prospective observational cohort study of children (<24 months) presenting to the emergency department at a single general hospital in Colombia with an acute respiratory infection with the aim of externally validating the reliability of the ReSVinet score [13]. The ReSVinet score was taken simultaneously by parents, paediatric doctor-professors (faculty) and residents in the emergency room. They were not blinded to the infant's outcome or to the other predictors. Information was gathered on enrolled infants via their electronic health records and phone consultation 10-days post-discharge. Outcomes analysed were hospitalisation and paediatric intensive care unit (PICU) admission.

The Hakizimana dataset was collected in a prospective cross-sectional study conducted in four tertiary hospitals (n=4) in Rwanda where infants (1-12 months) presenting with signs of respiratory distress were opportunistically enrolled to externally validate the ReSVinet score and the Liverpool Infant Bronchiolitis Severity Score (LIBSS) [5]. An anonymised version of this dataset was publicly available on the Harvard Dataverse [14]. The ReSVinet score of each infant was assessed independently by a nurse and resident within an hour of the infant presenting to

hospital. Further information on enrolled infants was gathered through standardised questionaries filled in by health-care providers involved in the delivery of care to the infant; as such they were not blinded to the infant's outcomes or to the other predictors. The dataset contained no missing values. Outcomes analysed were hospitalisations, admission to the high dependency unit (HDU), admission to the PICU, intubation or ventilation and death.

**Simplified score**

To develop the simplified scores, we only made use of the development dataset (RESCEU). Specifically we performed a grid search to explore all the possible combinations of parameters and identifed the equally-weighted combination of parameters giving the highest point estimate of area under the receiver operating characteristic (AUROC), for requirement of admission to hospital, admission to ICU, and need for SIMV.

After developing the candidate simplified scores, we assessed the AUROC for a range of outcomes in the external validation datasets (Hakizimana and Camacho-Cruz datasets).

The RESCEU dataset was used to develop the simplified score as it was larger (n=311 infants) compared to the and Camacho-Cruz (n=188) and Hakizimana (n=100) datasets which were used for validation.

**Validation of the original ReSVinet score**

We additionally used the RESCEU and Camacho-Cruz datasets to externally validate the discriminative validity of the original ReSVinet score by assessing the AUROC for the outcomes mentioned above. This was not done for the Hakizimana dataset as this has already been done and reported in the original publication [5].

**Statistical analysis**

For each of the above measures we calculated confidence intervals (CIs) of the AUROCs, and created p-value matrices. We considered a p-value <0.005 as constituting statistical significance, and used DeLong's method to calculate both CIs and paired p-values between AUROCs [15]. Paired p-values (as opposed to unpaired values) were calculated given that we were comparing AUROCs for the same outcomes in the same datasets. We used R version 4.1.2 and Python version 3.6.2 to conduct the analysis.

**Missing data**

In the development dataset, we excluded all infants with missing items for the ReSVinet score. The Hakizimana dataset had no missing data, as the participants with any missing data (n=7) had already been excluded in the dataset available publicly.

For outcome data, we excluded participants from specific analyses if the outcome was not available; in all cases we explicitly stated the number of participants included in each analysis.

## RESULTS

### Included participants

The RESCEU study enrolled 325 RSV cases of which 14 had missing data for the ReSVinet score parameters and so were excluded; 11 had no data for any of the score parameters and the remaining 3 had no data for two parameters each. As such, overall 311 patients were included. The Camacho-Cruz study enrolled 191 patients of which three were excluded as two did not meet the inclusion criteria and one was lost to follow up; as such, overall data on 188 patients were included. The Hakizimana study recruited 107 patients, of which for 7 data collection was incomplete and so were excluded in the publicly available dataset; overall data on 100 infants were included.

### Baseline characteristics

Baseline characteristic for included participants by dataset are summarised in *Supplementary Table 3*. Of note, the major difference across the datasets were the differences in age at enrolment across the 3 datasets. The median age at enrolment was 3, 7 and 10 months for the RESCEU, Hakizimana and Camacho-Cruz datasets respectively. Additionally across all 3 datasets there was a greater proportion of males to females.

### Outcome data

The outcomes for included participants by dataset are summarised in *Supplementary Table 4*. In the RESCEU dataset (n=311) 75% of the included patients were hospitalised, 28% admitted to the PICU and 25% who were ventilated. In the Hakizimana dataset (n=100) 93% of the included patients were hospitalised, 22% admitted to HDU, 7% admitted to PICU, 5% ventilated and 6% who died. In the Camacho-Cruz dataset (n=188) 3% were admitted to the ICU and there were no deaths.

### Development of simplified ReSVinet scores

Using all of the development dataset (i.e. the RESCEU dataset), the best combination of equally weighted parameters to maximise for the point-estimate of AUROC for discriminating by hospitalisation, ICU admission and SIMV requirement were determined; this produced three unique candidate simplified scores (ReSVinet-3, ReSVinet-4 and ReSVinet-6 – the suffix in this instance number of parameters included in the simplified score). ReSVinet-3 used the following parameters: respiratory difficulty, apnoea & general condition. ReSVinet-4 & ReSVinet-6 used the same parameters as ReSVinet-3 as well as medical intervention; and medical intervention, feeding intolerance and respiratory frequency respectively.

These scores, as assessed by their AUROC, were excellent in the RESCEU dataset at discriminating by the patients' hospitalisation status, ICU admission and requirement for SIMV (See [2,3] for cut-offs) (see *Table 1*). For ICU admission and requirement for SIMV, there were

no statistically significant differences in the performance of the original and candidate scores (see *Supplementary Table 5* for p-value matrices). For hospitalisations although there were no statistically significant differences between each candidate scores and the original score, there were significant differences between the candidate scores; specifically, ReSVinet-6 was marginally better than both ReSVinet-3 & ReSVinet-4, and ReSVinet-4 marginally better than ReSVinet-3.

We additionally externally validated the discriminative validity of the original ReSVinet score in the RESCEU dataset (see *Table 1*); in this dataset it was excellent at discriminating on hospitalisation, and borderline-excellent for ICU and need for SIMV.

**External validation**

Subsequently we externally validated the three candidate simplified scores using the external validation datasets (Camacho-Cruz & Hakizimana) (see *Supplementary Table 6*). In these datasets, they generally performed similarly to the original score. There were only two instances in which there were statistically significant differences (see *Supplementary Table 7* for the p-value matrices). The first occurred in the Hakizimana dataset where both ReSVinet-3 & ReSVinet-4 performed extremely poorly at discriminating for HDU admission but both the original ReSVinet score, and ReSVinet-6 performed moderately well. The second occurred in the Camacho-Cruz dataset where all of the candidate scores outperformed the original score, and ReSVinet-4 outperformed both ReSVinet-6 & ReSVinet-3.

We additionally externally validated the discriminative validity of the original ReSVinet score in the Camacho-Cruz dataset (see *Supplementary Table 6*); in this dataset it was moderate at discriminating for both hospitalization and PICU admission.

## DISCUSSION

Overall, the candidate simplified scores were moderate-excellent at discriminating, in both the development and external validation datasets, for a range of outcomes spanning the entire range of secondary-care disease severity. They performed equally well or better than the original ReSVinet score at discriminating by all of the outcomes evaluated; the only exception to this was in one dataset (Hakizimana) where two (ReSVinet-3 & ReSVinet-4) out of three of the candidate scores performed extremely poorly at differentiating by admission to the HDU. Additionally, the ReSVinet score was externally validated in two new datasets (Camacho-Cruz & Hakizimana) in which it performed moderate to excellent. Even though the RESCEU dataset (from infants in high income countries) was used to develop a simplified score to be applicable globally, we do not view this as a limitation. We validated the scores externally in datasets from low-income and middle-income countries ; this is relatively unusual as vast majority of previous validation efforts have occurred in high-income countries [2, 3].

In the three datasets used there were differences in when the ReSVinet score was taken. In the RESCEU dataset, the score was taken at the time of recruitment and so there is potentially a large degree of variability at the point in the disease course at which the scores were taken; whereas in the Hakizimana & Camacho-Cruz datasets the scores were taken within an hour of presentation and in the emergency room respectively. This greater variability reduces the usefulness of the validation in the RESCEU dataset. Additionally, it would be useful to have conducted additional analyses with a threshold on the time of the outcome since taking the score (e.g. discriminative validity of score at predicting need for admission to ICU within 24 hours) as the disease course is not always linear, and so we may be systematically making an error of the actual validity; this was not possible in our case as the required data for this was not collected.

There were also some differences in the inclusion criteria of the datasets, specifically the age ranges and, presence of RSV and comorbidities. The development dataset included 50 participants with comorbidities whereas the Camacho-Cruz dataset excluded all patients with significant comorbidities and the Hakizimana dataset excluded those with a history of chronic lung disease and those who presented with non-respiratory causes of respiratory distress. It is not clear, on the basis of our analysis, if these differences (age, RSV, comorbidities) affected the discriminative ability of the scores. However, recent analysis of the same published RESCEU dataset has indicated that fever is more commonly found in older infants (≥ 6 months), as well as more frequently in those with RSV-A and with higher viral loads [12]; however fever, viral load and RSV-subgroup were not found to be associated with severity.

To systematically assess bias, we used a modified version of the PROBAST tool to conduct an internal risk of bias assessment (see *Online supplementary file, appendix 2*). Based on this we deemed overall that there was a high risk that our estimates of the discriminative validity of these scores are biased.

The primary reason for this is due the small number of patients with positive outcomes owing to the small size of the datasets. Only in the RESCEU dataset for 1 outcome (hospitalisations) were the PROBAST thresholds (≥20 events per variable for validation [RESCEU]; ≥100 for validation [Hakizimana & Camacho-Cruz]) met. In 5 instances, the number of participants with the outcome of interests is a single-digit figure. This is similar to what is observed in published literature, presumably due to difficulties in collecting large datasets of these sorts. For illustration purposes, to be confident in reporting on the discriminative validity for death (assuming the same approximate ratio of deaths), the Hakizimana dataset would need to be approximately 17-times larger. Additionally, in an ideal world, we would have used multiple imputation for missing data instead of excluding patients with missing score-items and additionally assessed calibration; this was not done due to time constraints. Regardless, neither of these factors would have changed our overall classification of this study as having an associated high risk of bias.

To guide, as well as assess the transparency of reporting, the TRIPOD checklist was employed (see *Online supplementary file, appendix 1*). The overall score was 26/32 (81%). This is

considerably better than the average score, 53%, found in our systematic review [3], although there still is room for improvement. The issues with reporting were that not all required information was reported in the title (due to length constraints), the number of positive outcomes were not reported in the abstract (due to word limit constraints), the number of centres enrolment took place at for the RESCEU study was not reported, that interaction terms were not tested for and, that calibration was not reported (as not assessed).

**Implications on research, policy & practice**

The findings indicate that fever as a criterion is redundant and may be excluded from the original ReSVinet score (i.e. ReSVinet-6 should be adopted). However, given the small number of other validation studies (see *Supplementary Table 1*) and more generally studies evaluating the ReSVinet score (e.g. reliability, responsiveness, utility), we cannot at this time recommend use of either the original ReSVinet score or ReSVinet-6 in routine clinical care or as a primary outcome for clinical trials.

Further research on much-larger datasets is required to validate the usefulness of the original ReSVinet score, as well as the simplified scores we have developed. Given the relatively small number of patients with positive outcomes in the development dataset, it would also be appropriate to repeat this exercise, using similar methods, in larger datasets to examine if different combinations of parameters (i.e. scores) are selected. Ideally the datasets employed for these exercise as well as being larger, may be gathered in resource-limited settings where the burden is disproportionately high [2,3].

**References**


1. Justicia-Grande AJ, Pardo-Seco J, Cebey-López M, et al. Development and Validation of a New Clinical Scale for Infants with Acute Respiratory Infection: The ReSVinet Scale. PloS One [Internet]. **2016** [cited 2022 Mar 1]; 11(6):e0157665. Available from: https://pubmed.ncbi.nlm.nih.gov/27327497/
2. Hakizimana B, Saint G, Miert C van, Cartledge P. Can a Respiratory Severity Score Accurately Assess Respiratory Distress in Children with Bronchiolitis in a Resource-Limited Setting? Journal of Tropical Pediatrics [Internet]. **2020**; 66(2):234–243. Available from: https://academic.oup.com/tropej/article/66/2/234/5570307#:~:text=Though%20scores%20should%20measure%20what
3. Sheikh, Z, Potter E, Li Y, Cohen R, Santos GD, Bont L, Nair H. Validity of clinical severity scores for respiratory syncytial virus: a systematic review. Submitted to Journal of Infectious Diseases RSV-PROMISE supplement.
4. Justicia-Grande AJ, Martinón-Torres F. The ReSVinet Score for Bronchiolitis: A Scale for All Seasons. American Journal of Perinatology [Internet]. **2019** [cited 2022 Mar 13]; 36(S 02):S48–S53. Available from: https://www.thieme-connect.com/products/ejournals/pdf/10.1055/s-0039-1691800.pdf
5. Hakizimana B, Kalimba E, Ndatinya A, Saint G, Miert C van, Cartledge PT. Field testing two existing, standardized respiratory severity scores (LIBSS and ReSViNET) in infants presenting

with acute respiratory illness to tertiary hospitals in Rwanda – a validation and inter-rater reliability study. PLoS ONE [Internet]. **2021** [cited 2022 Mar 4]; 16(11):e0258882. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8568200/

6. Justicia-Grande, A.J., Pardo Seco, J., Alvarez Garnelo, M.E., Pena Nieto, J., Sánchez Lastres, J.M., Del Río Garma, M., Valls Duran, M.T., San José González, M.A., Regueiro Martínez, J., Alonso Alonso, C., Fernández Bustillo, J.M., García Sánchez, N., Castro Paz, L., López-Silvarrey Varela, Á., López Del Olmo, M.C., Rivero Calle, I., Obando Pacheco, P., Porto Silva, M., Grela Beiroa, A. and Serén, S. Validation of a Clinical Respiratory Score (the ReSVinet Scale) in Pediatric Patients Suffering from Respiratory Infection in a Primary Care Setting [Internet]. **2017**. Available from: https://espid2017.kenes.com/Documents/ESPID17%20abstracts.pdf
7. Tungsupreechameth A, Srisingh K. Factors Associated with Severe Lower Respiratory Tract Infection from Respiratory Syncytial Virus (RSV) in Thai Children. Siriraj Medical Journal. **2021**; 73(12):808–814.
8. Vandendijck Y, Scott J, Heerwegh D, et al. Measuring respiratory syncytial virus infection severity in hospitalized children using the Pediatric Respiratory Syncytial Virus Electronic Severity and Outcome Rating System (PRESORS). Influenza and Other Respiratory Viruses. **2022**;
9. Moons KGM, Wolff RF, Riley RD, et al. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. Annals of Internal Medicine. **2019**; 170(1):W1.
10. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. Annals of Internal Medicine. 2015; 162(1):W1.
11. Jefferies K, Drysdale SB, Robinson H, et al. Presumed Risk Factors and Biomarkers for Severe Respiratory Syncytial Virus Disease and Related Sequelae: Protocol for an Observational Multicenter, Case-Control Study From the Respiratory Syncytial Virus Consortium in Europe (RESCEU). The Journal of Infectious Diseases. **2020**; 222(Supplement_7):S658–S665.
12. McGinley JP, Lin GL, Öner D, et al. Clinical and Viral Factors Associated With Disease Severity and Subsequent Wheezing in Infants With Respiratory Syncytial Virus Infection. The Journal of Infectious Diseases. **2022**; 226(Supplement_1):S45–S54.
13. Camacho-Cruz J, Briñez S, Alvarez J, Leal V, Gómez LV, Vasquez-Hoyos P. Use of the ReSVinet Scale for parents and healthcare workers in a paediatric emergency service: a prospective study. BMJ Paediatrics Open [Internet]. **2021** [cited 2022 Apr 22]; 5(1):e000966. Available from: https://bmjpaedsopen.bmj.com/content/5/1/e000966
14. Cartledge P. LIBBS and ReSVinet Rwanda - Data (V9 for DOI).csv, Field testing two existing standardised respiratory severity scores (LIBSS and ReSVinet) in infants presenting with acute respiratory illness to tertiary hospitals in Rwanda [Internet]. Harvard Dataverse. **2020** [cited 2022 Aug 26]. Available from: https://dataverse.harvard.edu/file.xhtml?persistentId=doi:10.7910/DVN/N4O05G/AKSNOZ
15. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. Biometrics. **1988**; 44(3):837.

**PROMISE Investigators**

Harish Nair, Harry Campbell, Richard Osei-Yeboah (University of Edinburgh); John Paget (NIVEL); Philippe Beutels (Universiteit Antwerpen); Anne Teirlinck (RIVM); Hanna Nohynek (THL); Louis Bont (Univeristy Medical Center Utrecht); Andrew Pollard (University of Oxford); Peter Openshaw (Imperial College London); You Li (Nanjing Medical University); Jeroen Aerssens, Gabriela Ispas (Janssen); Veena Kumar (Novavax); Tin Tin Htar, Elizabeth Begier, Jessica Atwell (Pfizer); Charlotte Vernhes, Rolf Kramer, Mathieu Bangert (Sanofi Pasteur); Gaël Dos Santos, Rachel Cohen, Theo Last (GSK); Bahar Ahani (AstraZeneca); Nuria Machin (TeamIT)

***Author contributions:*** HN conceived the idea. HN, YL & ZS designed the approach. SBD, HR, JGW, JM, GLL, DÖ, JA, AJG, FMT, AJP, and LB collected and contributed data. ZS conducted the analysis and authored the manuscript. DÖ, JA, AJG, FMT, EP, YL, LB & HN commented critically on several drafts of the manuscript. PROMISE investigators reviewed the manuscript prior to submission.

***Conflict of interest statement:*** None for ZS, EP, JM, GLL or HR.

YL has received funding from Wellcome Trust and GSK outside the submitted work. YL received personal fees from Pfizer.

SBD has received honoraria from MSD and Sanofi for taking part in RSV advisory boards and has provided consultancy and/or investigator roles in relation to product development for Janssen, AstraZeneca, Pfizer, Valneva, MSD, iLiAD and Sanofi with fees paid to St George's, University of London.

JGW participated in the advisory board of Janssen with fees paid to UMCU.

DÖ and JA are employees of Janssen Pharmaceutical NV (Beerse, Belgium).

AJG does not (and did not) receive any form of payment from private institutions. AJG is not the subject of any grant (including public ones) and he is currently not involved in clinical trials concerning the area covered by this document.

FM-T reports grants from Janssen, MSD, GSK, and Astra Zeneca, personal fees from Ablynx, GSK, Pfizer, MSD, Moderna, Astra Zeneca, Sanofi-Pasteur, Novavax, Seqirus and Biofabri, non-financial support from GSK, Pfizer, MSD and Seqirus, and trial fees paid to his institution from these different companies (except Biofabri).

AJP is Chair of the UK Department of Health and Social Care's Joint Committee on Vaccination and Immunisation. The University of Oxford has entered into a partnership with AstraZeneca for the development of a COVID19 vaccine.

LB has received funding through UMC Utrecht from AbbVie, Janssen, the Bill and Melinda Gates Foundation, Nutricia Danon, MeMed Diagnostics, GSK, , AstraZeneca, Sanofi, Ablynx, Bavaria Nordic, MabXience, Novavax and Pfizer.

HN has received funding from IMI, NIHR and Pfizer. HN received fees from GSK, AbbVie, AZ, Novavax, Sanofi, Merck and ReViral.

***Funding statement:*** This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 101034339. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA. This publication only reflects the author's view and the JU is not responsible for any use that may be made of the information it contains herein.

Mention of any meeting(s) where the information has previously been presented

Presented at PROMISE 3rd General Annual Meeting (September, 2022).

Poster presented at the 12th International RSV Symposium (October, 2022).

***Corresponding author contact information***

Professor Harish Nair

Usher Institute

University of Edinburgh

Old Medical School

Teviot Place

Edinburgh

EH8 9AG, UK

Harish.Nair@ed.ac.uk

**Table 1:** AUROC for original ReSVinet & candidate simplified scores in development (RESCEU) dataset (n=311)

| **Maximising AUROC for** | **Name** | **Parameters included** | **Parameters excluded** | **AUROC (95% CI)** | | |
|---|---|---|---|---|---|---|
| | | | | *Hospitalisation* | *ICU Admission* | *SIMV requirement (n=280)* |
| | ReSVinet | Feeding intolerance Medical intervention<br>Respiratory difficulty<br>Respiratory frequency<br>Apnoea<br>General condition<br>Fever | | 0.9437 (0.9191-0.9683) | 0.8995 (0.8629-0.9361) | 0.8958 (0.8551-0.9364) |
| **Hospitalisation** | ReSVinet-6 | Feeding intolerance Medical intervention<br>Respiratory difficulty<br>Respiratory frequency<br>Apnoea<br>General condition | Fever | 0.9484 (0.9257-0.9711) | 0.9080 (0.8734-0.9427) | 0.9069 (0.8684-0.9454) |
| **ICU Admission** | ReSVinet-3 | Respiratory difficulty<br>Apnoea | Feeding intolerance<br>Medical | 0.9165 (0.8867-0.9464) | 0.9156 (0.8836-0.9477) | 0.9164 (0.8819-0.951) |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | General condition | intervention<br>Respiratory frequency<br>Fever | | | |
| **SIMV requirement (n=280)** | ReSVinet-4 | Medical intervention<br>Respiratory difficulty<br>Apnoea<br>General condition | Feeding intolerance<br>Respiratory frequency<br>Fever | 0.9378 (0.9135-0.9621) | 0.9150 (0.8818-0.9483) | 0.9234 (0.8892-0.9576) |

(AUROC- Area under the receiver-operator characteristic curve; CI - confidence interval; ICU - intensive care unit; RESCEU- REspiratory Syncytial virus Consortium in Europe; SIMV - synchronised intermittent mandatory ventilation)