# *The predictive value of the Apgar score for morbidity and mortality for newborns born to women with hypertensive disorders during pregnancy in Ghana:* *a prospective cohort study.*

| | |
|---|---|
| Name student | Eline E.M.Y. Pelders |
| Student number | 6361668 |
| Thesis version | Final version |
| Date | June 24 2021 |
| Supervisor | Klaartje Olde Loohuis & Joyce Browne |
| Lecturer | Jacqueline van Dijk |
| Internship institution | Julius Centrum, UMC Utrecht |
| Journal | PLOS one |
| Master | Clinical Health science, Nursing science |
| Institution | University Utrecht, UMC Utrecht |
| Number of words summary English version | 298 |
| Number of words summery Dutch version | 289 |
| Number of words full-text | 3747 |
| STROBE criteria are used for the purpose of transparent reporting | |

Abstract

Background: A child born in sub-Saharan Africa is ten times more likely to die in the first month than a child born in a high-income country. The Apgar-score was developed to reduce infant mortality and morbidity through early detection of complications. Despite its widespread use, the accuracy of the predictive value of the Apgar-score for morbidity and mortality in low- and middle-income countries has scarcely been evaluated. None focused on newborns born to women with severe hypertensive disorders during pregnancy, while this is a major reason for neonatal mortality and morbidity.

Aim: To determine the predictive value of the five-minute Apgar-score for morbidity and mortality in newborns born to women who experienced severe hypertensive disorders during pregnancy and were admitted to referral hospitals in Ghana.

Method: A secondary data analysis was performed on data collected between December 2017 and March 2021 within a prospective cohort study (SPOT study) in six major obstetric care centres in Ghana. The study population consists of newborns born to women who experienced hypertensive disorders during pregnancy in Ghana. Multivariable logistic regression analyses and a receiver operating characteristic (ROC) curve were used to determine the predictive value of the five-minute Apgar-score.

Results: In total, 565 newborns were included. Univariate logistic regression analyses found an association between Apgar-score and all neonatal adverse events (OR between 1.32-2.23). In the multivariate logistic regression analysis only NICU-admission lost the association with Apgar-score. The ROC (univariate) indicated an acceptable to good discriminative ability (AUC between 0.67-0.78) for all neonatal adverse outcomes. The multivariate AUC slightly increased compared to the univariate model (AUC between 0.72-0.79).

Conclusion and Recommendations: The five-minute Apgar-score could be used as a warning sign for all adverse outcomes. Future research is needed to refine the prediction model further.

Keywords: "Apgar score", “morbidity”, “mortality”, “pre-eclampsia”, “Africa”

## Samenvatting

Achtergrond: Een kind geboren in sub-Sahara Afrika heeft tien keer meer kans om te overlijden in de eerste levensmaand dan een kind geboren in een land met modaal inkomen. De Apgar-score is ontwikkeld om sterfte en ziekte bij zuigelingen te verlagen door vroegtijdig complicaties te signaleren. Ondanks het vele gebruik is de voorspellende waarde van de Apgar-score voor morbiditeit en mortaliteit in lageloonlanden zeer beperkt onderzocht. Verder is er geen enkele studie over de voorspellende waarde van de Apgar-score in sub-Sahara Afrika die specifiek gefocust was op pasgeborenen van vrouwen met een hypertensieve aandoening tijdens de zwangerschap, terwijl dit wel een belangrijke oorzak van neonatale ziekte en sterfte is.

Doel: De voorspellende waarde van de vijf-minuten Apgar-score bepalen voor ernstige morbiditeit en mortaliteit bij pasgeborenen van vrouwen met een hypertensieve aandoening tijdens de zwangerschap die waren opgenomen in Ghana.

Methoden: Een secondaire data-analyse is uitgevoerd met data verzameld tussen December 2017 en Maart 2021 door een prospectieve cohortstudie (SPOT studie) in zes grote verloskundige zorgcentrums in Ghana. De studiepopulatie bestaat uit pasgeborenen van vrouwen met een hypertensieve aandoening tijdens de zwangerschap in Ghana. (Multipele) logistische regressieanalyse en de ROC-curves (Engels: Receiver Operating Characteristic) zijn gebruikt om de voorspellende waarde van de Apgar-score te bepalen.

Resultaten: In totaal zijn er 565 pasgeborenen geïncludeerd. Middels logistische regressie werd een associatie aangetoond tussen de Apgar-score en morbiditeit/mortaliteit (OR tussen 1.32-2.23). Het multipele model toonde geen associatie meer ten opzichte van NICU-opname. De ROC toonde een acceptabel tot goed onderscheidend vermogen aan voor Apgar-score. In het multivariate model verbeterde het onderscheidend vermogen (AUC tussen 0.72-0.79).

Conclusie: De vijf-minuten Apgar-score kan gebruikt worden als waarschuwing voor mortaliteit en morbiditeit. Vervolgonderzoek is nodig om het predictiemodel te verbeteren.


Sleutelwoorden: “Apgar-score”, “morbiditeit”, “mortaliteit”, “pre-eclampsie”, “Afrika”

## Introduction

Globally, the infant mortality rate has significantly decreased since 1990[1]. However, in 2017, 4.1 million infants died worldwide[1]. According to the World Health Organization (WHO), the risk of an infant dying is the highest in sub-Saharan Africa; over six times higher than the risk in Europe[1]. The neonatal period, the first 28 days after birth, represents the most vulnerable time for a child's survival[1]. In 2016, an estimated forty-six per cent of all infant deaths occurred during the neonatal perdiod[2]. Globally, substantial progress has been made in child survival since 1990[2]. However, the decline rate of neonatal mortality has been slower than that of the infant mortality rate[2]. A child born in sub-Saharan Africa is ten times more likely to die in the first month than a child born in a high-income country (HIC)[2].

Given these statistics, contributing factors were investigated for the high neonatal mortality rate in low- and middle-income countries (LMICs)[3,4]. These reports indicate that substandard care, inadequate training, low staff competence and a lack of resources, including equipment and medication, are factors that contribute to neonatal deaths[3,4]. Also, there is evidence for an association between low nurse staffing levels and patients' outcomes, most notedly patient mortality[5]. Therefore, a lack of healthcare coverage is likely to affect neonatal outcomes and could contribute to neonatal deaths[6].

Additionally, sub-Sahara Africa had the highest rates of preterm birth in contrast to Europe in 2005[7]. Preterm birth is defined as childbirth occurring at less than 37 completed weeks or 259 days of gestation[7]. Preterm birth is especially common when the mother has a hypertensive disorder during pregnancy (HDP), like pre-eclampsia or eclampsia[8]. Hypertensive disorders are the most common complication during pregnancy, occurring in five to eleven per cent of pregnancies[9] and is a major reason for neonatal mortality and morbidity[7]. A report on LMICs in Africa stated that more than fifty per cent of preterm neonatal deaths can be prevented when neonatal units have adequate staffing levels, proper follow-up, early detection and timely management of complications[10].

In order to prevent neonatal deaths and early detection of complications, Virginia Apgar developed a prediction score in 1952[11,12]. This score evaluates a newborns clinical status rapidly after birth to assess the need for immediate medical intervention[11,12]. Almost 70 years later, the Apgar score (AS) is globally used to measure the health status of newborns based on heart rate, respiration, colour, muscle tone and reflex irritability[12]. A score (zero, one or two) is assigned to each parameter, and the sum of the scores is the AS[13]. The AS is categorised into three levels[13]: low (0-3), intermediate (4-6) and normal to excellent (7 or higher)[12]. Since the incidence of an Apgar score in the range of 0-3 is low, studies often combine low and intermediate scores to examine outcomes among infants with a score from 0-7[13].

Besides measuring the newborns health status rapidly after birth, the AS is also used as a prognostic tool for infant mortality and severe morbidity[11]. At first, the AS was only taken at the first minute after birth[11]. Later, the five minute AS was added, as it is a better predictor for neonatal survival[11]. Several large studies found an association between a low AS and increased risk of neonatal and infant mortality in both preterm and term newborns[14–16], but this has not been extensively evaluated in LMICs contexts[17–19]. It is necessary to evaluate this since evidence-based decision-making on appropriate treatment choices is critical to reduce deaths and appropriate use of limited healthcare resources in LMICs[20]. In addition to limited evaluation, none of the LMIC-based studies determined whether the predictive value of the AS, on morbidity or severe mortality, is different for newborns born to women who experienced severe hypertensive disorders in early gestation (26-34 weeks). For proper risk selection and initiation of appropriate care and treatment, it is crucial to be able to predict which infants have a higher risk of morbidity and mortality[20]. Therefore, this study aims to investigate the predictive performance of the AS for adverse neonatal outcomes for women with HDP.

## Aim

This study aimed to determine the predictive value of the five-minute AS for severe morbidity and mortality (within six weeks of birth) for newborns born to women who experienced hypertensive disorders in early gestation (26-34 weeks) and were admitted to referral hospitals in Ghana.

## Method

### Design

The current study (prospective cohort study) used the dataset of the Severe Pre-eclampsia adverse Outcome Triage (SPOT) study[21], a prospective cohort study in Ghana. The SPOT study aims to improve the quality of care for women with HDP in LMICs. Because of the shared interest, the SPOT study fits the purpose of the current study perfectly, and it was appropriate and efficient to do a secondary data analysis on their data[22]. The current study uses the SPOT study data collected between December 2017 and March 2021. The study is reported according to the Reporting of Observation Studies in Epidemiology (STROBE) statement[23].

### Population and domain

During data collection, the SPOT study collected 603 births from women with HDP admitted between 26 and 34 weeks of gestation to one of the six major obstetric care centres in the Greater Accra Region of Ghana. All newborns born from the participating women from the SPOT study were included for the current study. Stillbirths were excluded from participating in the current study. Since, the AS is only measured for live births.

**Sample size.** No sample size calculation was conducted before the start of the current study. Instead, a pragmatic approach was used with a theoretical foundation to choose the sample size. On the one hand, the current study was dependent on the available data of the SPOT study. On the other hand, a minimum of at least 100, but preferably 200 or more, adverse events is recommended for externally validating a prognostic model[24]. Given the available data and the recommendations a minimal of 100 adverse events per outcome was aspired and preferably 200 or more[24]. Therefore, all available samples from the SPOT study were used.

### SPOT data collection procedure

The SPOT study recruited all pregnant women with HDP who met the eligibility criteria. Consenting women had a baseline assessment within 24 hours after admission. Data were collected prospectively by trained local staff (midwives or physicians), using standardised data extraction forms and a case report form. The variables for the current study were collected from medical records since these variables were measured during routine workup at admission to and during the stay at the hospital. During a routine visit six weeks after delivery, maternal and infant routine check-ups took place, and the follow-up data were also derived from the medical record.

### Current study variables

**Baseline characteristics**. This study selected five baseline characteristics for the newborns and eight for the women. Baseline characteristics for the newborns included AS

(high Apgar (≥7)/low Apgar (<7)), sex (female/male), gestational age (GA) at delivery (total weeks), birth weight (grams) and congenital anomalies (yes/no). Baseline characteristics for the women were maternal age (years), ethnic origin (Akan/Ewe/Ga-Dagma/Northem/Hausa/other), religion (Christian/Muslim/Traditional/other), educational status (No formal education/Primary/JSS/SSS/Tertiary/other), marital status (married/single/cohabiting/widowed/other), nicotine use during pregnancy (yes/no), alcohol use during pregnancy (yes/no) and drug use during pregnancy (yes/no).

**Main outcomes of interest**. Neonatal adverse outcomes were divided into two groups: morbidity and mortality. Morbidity is expressed in resuscitation/ventilation >1 min (yes/no), NICU admission (yes/no) and serious neonatal morbidities within six weeks(yes/no). Mortality is a composite outcome derived from two original variables: survival during hospital admission (yes/no) and death within six weeks (yes/no). The follow-up data (serious neonatal morbidities within six weeks and death within six weeks) was measured during a routine maternal and infant check-up six weeks after delivery.

**Data analysis**

The raw obtained data file was cleaned to prepare the data for analysis[25]. Additionally, the data were checked for missing values[25]. Due to the quantity of the missing values and the cut-off value of ten per cent missings per variable was exceeded for several characteristics, complete case analysis was not feasible[26]. The nature of the missing data was assessed using a correlation matrix (heatmap)[25]. The Multivariate Imputation by Chained Equations (MICE) method was appropriate for imputing missing values for the baseline characteristics of the newborns and women to prevent biased statistical analysis[27,28]. Missing values for the adverse neonatal outcomes and the AS were not imputed. The imputation model was checked and refined by assessing summary statistics (proportions, means and standard deviations) that provided information on the observed values, imputed values, and the combined values[27]. Also, the number of iterations was determined based on the stability of the imputation model[27].

The data were analysed using software programs Python (version 3.8) and R (version 4.0). The following software packages were used for Python: Pandas (version 1.1), Missingno (version 0.4), Numpy (version 1.19), scipy.stats (version 0.12) and scipy (version 1.5). For R also software packages were used: 0.7.4, Caret (version 6.0), Epitools (version 0.5), PROC (version 1.17) and Tidyr (version1.1). In preparation for the primary analysis, descriptive analyses of the newborns/women baseline characteristics and outcomes were performed. The descriptive analyses were used to gain insight into the distribution of the variables. For categorical variables, absolute (n) and proportions (%) were calculated. For the continuous variables, means and standard deviations (SD) were calculated. Descriptive

analysis was performed to to gain insight into both high and low Apgar groups. In addition to the descriptive analyses, the two groups were compared with a Chi-squared test (categorical data) or an independent samples t-test (continuous data).

A univariate logistic regression analysis was performed to compare newborn'/women's characteristics and adverse neonatal outcomes separately. This was used to calculate the odds ratio (OR) and 95% confidence interval (CI).

Potential confounders were selected based on clinical and epidemiological knowledge, literature[16] and were partly data-driven. GA was selected in advance as a confounder and included in the models to evaluate the association controlling for GA using a multiple logistic regression analysis (adjusted OR). In addition to GA, significant interaction terms (tested between GA*birth weight and every neonatal adverse outcome separate with logistic regression) and significant values from the univariate analysis were also added in the multiple logistic regression model. A *P*-value below 0.05 was considered statistically significant.

To estimate the prediction of neonatal adverse outcomes using AS, receiver operating characteristic (ROC) curves were plotted based on univariate (AS as only predictor) and multivariate (AS with other significant variables as predictor) models. Based on the ROC curve, the area under the curve (AUC) was calculated[29,30].

In addition to the AUC score, sensitivity (SENS), specificity (SPEC), positive predictive value (PPV), and negative predictive value (NPV) were calculated based on a confusion matrix[31]. For these values, AS was used as only predictor and needed to be binary. Therefore, the AS was divided into high and low Apgar groups.

**Ethical issues**

This study was conducted according to the principles of the Declaration of Helsinki (October 2013, 64$^{th}$ version)[32]. The researcher received approval from the Ghana Health Service Ethics Review Committee (GHS-ERC) and is GCP (Good Clinical Practice) certified. Permission or informed consent (IC) of the included women and newborns was not needed due to the comparable overall aim of the current study and the SPOT study. Therefore, the SPOT IC was sufficient for the current study. Also, the current study used a pseudonymised dataset, and linking tables were only available to the principal investigators of the SPOT study.

## Results

**Baseline characteristics**. In total, the SPOT study included 603 women between December 2017 and March 2021. After excluding stillbirths and adding the multiple births, 565 newborns were included in the current study (Figure 1). Of these newborns (without missings), 192 (47.3%) were male, and 214 (52.7%) were female. With a total mean birth weight of 1840 ± 832 gram and a GA of 33.3 ± 3.4 weeks (Table 1). Of all newborns, 150 (26.5%) had low AS at five minutes. The low Apgar group weighed 1265 ± 448 grams at birth and had a GA of 31.0 ± 2.6 weeks; values were significantly different from the high Apgar group (both *P*-values 0.000), high Apgar newborns were heavier on average (2098 ± 838 grams) and had a longer GA (34.5 ± 3.1 weeks). Of the 150 newborns with low AS, 54 (36.0%) died, and 17 (11.3%) got severe neonatal morbidities within six weeks. In high Apgar group, 35 (10,3%) newborns died, and 9 (2.6%) got servere neonatal morbidities within six weeks, which was significantly different between the two groups (mortality *P*-value 0.000/morbidity *P*-value 0.002). In addition to 'mortality' and 'morbidity', NICU admission (*P*-value 0.000) and resuscitation/ventilation >1 min (p-value 0.000) also were significantly different between the two groups. The characteristics of the women with low/high AS were not statistically different between the two groups.

<Figure 1 and Table 1>

**Missing values**. The distribution of the missing values is displayed in Table 2. The percentage of missing values per variable was between 2.3-35.0%, and the AS specificly contained 13.1% missing values. The majority of the missing data was Missing At Random (MAR)[25], except for the variables from the six-week follow-up appointment. Remarkable is that these missing values (morbidity six weeks and mortality six weeks) are largely missing from the high Apgar group, which indicates that these missing variables are connected to the AS of a newborn and therefore are Missing Not At Random (MNAR)[25].

<Table 2>

**Univariate and multivariate analysis**. Table 3 presents the univariate and multivariate logistic regression analyses for each outcome separately. The imputed dataset was used for these analyses and Figure 1 shows the number of cases used per adverse outcome. According to the univariate logistic regression analysis for 'NICU admission', AS (OR 2.23 (Standard Error (SE) 0.12)), sex (OR 0.40 (SE 0.22)), GA (OR 1.93 (SE 0.06)), birth weight (OR 1.00 (SE 0.0002)) and maternal age (OR 1.07 (SE 0.02)) are independently associated. For 'Resuscitation/ventilation', AS (OR 1.55 (SE 0.68)), GA (OR 1.20 (SE 0.039)), birth weight (OR 1.00 (SE 0.0002)) and maternal age (OR 1.06 (SE 0.02)) are independently associated. For 'morbidities within six weeks', AS (OR 1.32 (SE 0.10)), and GA (OR 1.25 (SE 0.07)), are independently associated. For 'mortality within six weeks', AS

(OR 1.49 (SE 0.07)), GA (OR 1.36 (SE 0.05)), and birth weight (OR 1.00 (SE 0.0003)), are independently associated.

It was necessary to adjust for potential confounders. Therefore, the statistically significant variables from the univariate analysis, the previous selected confounder GA and significant interactiterms (GA*birth weight) were added to every multivariable logistic regression model. Multivariable logistic regression analysis for 'Resuscitation/ventilation' (AOR 1.42 (SE 0.078)), 'morbidities within six weeks' (AOR 1.42 (SE 0.11)) and 'mortality within six weeks' (AOR 1.31 (SE 0.08)) AS was associated. 'NICU admission' was no longer significantly associated with AS in the multivariate model.

For the variables congenital anomalies, drug and alcohol use during pregnancy, no logistic regression analysis was performed due to the low occurrence of these variables and recommendations for a minimal event per variable (EPV) rate of 10 in order to conduct logistic regression analyses[33,34].

<Table 3>

**Prediction of neonatal adverse outcomes using Apgar.** Table 4 presents the NPV and PPV, the sensitivity and specificity values (with AS as predictor) and the AUC scores. If these values came close to one (or 100%), it would indicate a perfect prognostic model[31]. Only NPV for morbidity within six weeks had a score close to one (0.86). However, PPV for morbidity within six weeks had a score of 0.05. No values were close to one for sensitivity and specificity for any adverse outcome.

ROC curves were plotted (Figure 2) and AUC calculated. The AUC for the univariate model indicated that the AS had an acceptable to good discriminative ability with AUC scores between 0.67-0.78. The AUC scores for the multivariate model indicated a slight increase, compared to the de model with AS as the only predictor, with scores between 0.72-0.79. Since 'NICU admission' was no longer associated in the multivariate model with AS, no ROC curve was made, and the AUC was not calculated for this variable.

<Figure 2 and Table 4>

## Discussion

The current study aimed to determine whether the five-minute AS has a predictive value for severe morbidity and mortality (within six weeks of birth) for newborns born to women who experienced hypertensive disorders in early gestation (26-34 weeks) and were admitted to referral hospitals in Ghana.

**Principal findings.** A significant difference for all the neonatal adverse outcomes (resuscitation/ventilation, NICU-admission, morbidity and mortality within six weeks) was found between the newborns with low and high AS. Also, univariate analyses indicated that there were significant associations between AS and all these neonatal adverse outcomes. Besides the univariate model, the AUC derived from the ROC curve indicated that the AS had an acceptable to good discriminative ability for all neonatal adverse outcomes when the univariate analysis was used. The AUC from the multivariate model slightly increased compared to the univariate model. The slight increase was expected because there were more variables to predict an adverse event with. Despite the acceptable to good discriminate ability, sensitivity, specificity, NPV, and PPV almost all lacked performance. A possible reason is that the AS was made binary and therefore lost power.

**Comparison with other studies.** The significant association found between AS and mortality is comparable with the population-based cohort study by Iliodromiti et al.[16]. They reported an powerful association between low Apgar at five minutes and the risk of neonatal death in Scotland between 1992 and 2010 (relative risk (RR) of 359.4 (95% CI 277.3-465.9) for early neonatal death and RR of 30.5 (95% CI 18.0–51.6) for late neonatal death). Their findings showed that the AS, continues to be prognostic for neonatal mortality. Also, Cnattingius et al.[35], and Lee et al.[36], indicated that the five minutes AS is a predictor for infant mortality. Lee et al.[36] also support the continued use of the AS as a predictor variable in preterm infant research and risk adjustments as well. Moreover, Cnattingius et al.[35] found that the five minutes AS may be a better predictor of neonatal mortality in very preterm infants and indicated that a low AS is a severe warning sign in term infants. However, their findings can only be generalised to other populations with similar baseline and ethnic characteristics which does not include Ghana, since LMICs, such as Ghana, are more likely to have more comorbidity, lower socioeconomic status, a lack of available resources, and differences in disease management compared to HIC[3,4,37]. These factors might alter the effect of the predictors on the outcome[37]. In contrast to the findings of previously discussed studies, the American Academy of Paediatrics Committee on Fetus and Newborn (AAPCFN) and the American College of Obstetricians and Gynaecologists Committee on Obstetrical Practice (ACOGCOP) expressed that the AS does not predict individual neonatal mortality

and should not be used for that purpose[38]. However, this study is also not similar in baseline and ethnic characteristics to be able to compare results with the current study. Eventually, one study was found with similar baseline and ethnic characteristics[17]. This study conducted in Uganda found that the AS is a helpful tool for predicting immediate outcomes (first 48 hours)[17]. Unfortunatly, no knowledge about the predictive performance of the AS within six weeks was investigated[17].

**Strengths and limitations.** The current study is the first to investigate whether the five-minutes AS could predict neonatal outcomes for women with HDP in a LMIC and could therefore be essential for clinical practice in countries like Ghana. In addition to the clinical importance, a strength is the secondary data analysis[22]. Typically, performing research inevitably involves some degree of risk for the participant. However, secondary data analysis has the benefit of answering the research question without putting, in this case, the newborns/women at additional risks[22]. Besides this benefit, the research question can be answered within less time and with lower costs than other research approaches[22]. In this case, it is especially relevant since to answer the current research question, many subjects and a follow-up of six weeks were required. Another strength was the use of multiple imputations to deal with missing values. As opposed to single imputations, multiple imputations account for the statistical uncertainty in the imputations preventing biased statistical analysis[27,28].

A few limitations of this study need to be addressed. One of the limitations was the used methods of the SPOT study, which was initially planned for a different research question. Thus, some used methods and the collected variables are inevitably different from those that otherwise would have been selected[39]. For example, the six-week follow-up data for morbidity and mortality. Preferably, a follow-up at 28 days after birth (neonatal period) was added. Since, this represents the most vulnerable time for a child's survival[1] and to be able to be more specific about the time of an adverse outcome happening. Another limitation is the high proportion of missing values. A possible explanation for this is that the current study analysis was done while data quality and completion efforts were still ongoing within the SPOT study. The missing values from the six-week follow-up (mortality and morbidity) are especially striking. Since, a large proportion of the missing values were within the high Apgar group. Therefore, these missing values were assessed as being MNAR. An possible explanation for these missing values is that mothers did not feel it necessary to come to the routine check-up six weeks after birth because they assessed their newborns as being healthy. Unfortunately, this cannot be confirmed by any other measured variable, so it remains a speculation. In conclusion, these outcomes may produce biased estimates, which should be considerd when looking at the results[40].

**Implications for clinical practice and future research**. The AS was developed to provide an early assessment of neonatal conditions. The current study shows that the five-minutes AS also could predict some adverse outcomes within six weeks of birth. However, these results need to be interpreted with caution. Clinical practice in LMICs should consider a low AS as a warning sign in newborns born from women with HDP for developing neonatal adverse outcomes and be aware of the potentional additional risk that these newborns have for developing adverse outcomes. Future research is needed to establish the predictive value of the five-minutes AS further within this population and indicate more predictors that could refine the prediction model for adverse outcomes within six weeks of birth. Additionally, breakdowns of the various components of the AS (heart rate, respiration, colour, muscle tone and reflex irritability) could be helpful to see which component may be the most clinically relevant in predicting adverse outcomes for these newborns specific.

References

1. WHO. Child Mortality and Causes of Death [Internet]. 2013 [cited 2020 Nov 7]. Available from: https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/child-mortality-and-causes-of-death

2. Newborns: improving survival and well-being [Internet]. World health Organisation. 2020 [cited 2020 Sep 3]. Available from: https://www.who.int/news-room/fact-sheets/detail/newborns-reducing-mortality

3. Korejo R, Bhutta S, Noorani KJ, Bhutta ZA. An audit and trends of perinatal mortality at the Jinnah Postgraduate Medical Centre, Karachi. Parity. 2007;31(40):40.

4. Hasan IJ, Khanum A. Health care utilization during terminal child illness in squatter settlements of Karachi. J Pakistan Med Assoc. 2000;50(12):405.

5. Griffiths P, Recio-Saucedo A, Dall'Ora C, Briggs J, Maruotti A, Meredith P, et al. The association between nurse staffing and omissions in nursing care: a systematic review. J Adv Nurs. 2018;74(7):1474–87.

6. Ehrenstein V. Association of Apgar scores with death and neurologic disability. Clin Epidemiol. 2009;1:45.

7. Beck S, Wojdyla D, Say L, Betran AP, Merialdi M, Requejo JH, et al. The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity. Bull World Health Organ. 2010;88:31–8.

8. JD GRLCJFI, Romero R. Epidemiology and causes of preterm birth. Lancet. 2008;371(9606):75–84.

9. Rivera-Romero O, Olmo A, Muñoz R, Stiefel P, Miranda ML, Beltrán LM. Mobile health solutions for hypertensive disorders in pregnancy: scoping literature review. JMIR mHealth uHealth. 2018;6(5):e130.

10. Tamene A, Abeje G, Addis Z. Survival and associated factors of mortality of preterm neonates admitted to Felege Hiwot specialized hospital, Bahir Dar, Ethiopia. SAGE open Med. 2020;8:2050312120953646.

11. Li F, Wu T, Lei X, Zhang H, Mao M, Zhang J. The Apgar Score and Infant Mortality. PLoS One [Internet]. 2013 Jul 29;8(7):e69072. Available from: https://doi.org/10.1371/journal.pone.0069072

12. Apgar V. A proposal for a new method of evaluation of the newborn. Class Pap Crit

Care. 1952;32(449):97.

13. Jehan I, Harris H, Salat S, Zeb A, Mobeen N, Pasha O, et al. Neonatal mortality, risk factors and causes: A prospective population-based cohort study in urban Pakistan. Bull World Health Organ. 2009 Feb;87(2):130–8.

14. Drage JS, Kennedy C, Schwarz BK. The Apgar score as an index of neonatal mortality: a report from the Collaborative Study of Cerebral Palsy. Obstet Gynecol. 1964;24(2):222–30.

15. Casey BM, McIntire DD, Leveno KJ. The continuing value of the Apgar score for the assessment of newborn infants. N Engl J Med. 2001;344(7):467–71.

16. Iliodromiti S, Mackay DF, Smith GCS, Pell JP, Nelson SM. Apgar score and the risk of cause-specific infant mortality: a population-based cohort study. Lancet. 2014;384(9956):1749–55.

17. Ondoa-Onama C, Tumwine JK. Immediate outcome of babies with low Apgar score in Mulago Hospital, Uganda. East Afr Med J. 2003;80(1):22–9.

18. Paul VK, Singh M, Sundaram KR, Deorari AK. Correlates of mortality among hospital-born neonates with birth asphyxia. Natl Med J India. 1997;10:54–6.

19. Nathoo KJ, Chimbira THK, Mtimavalye LAR. Mortality and immediate morbidity in term babies with low Apgar scores (Zimbabwe). Ann Trop Paediatr. 1990;10(3):239–44.

20. Houweling TAJ, van Klaveren D, Das S, Azad K, Tripathy P, Manandhar D, et al. A prediction model for neonatal mortality in low-and middle-income countries: an analysis of data from population surveillance sites in India, Nepal and Bangladesh. Int J Epidemiol. 2019;48(1):186–98.

21. Browne J, Sronfeyoh E, Klipstein-Grobusch K, Geest I van der, Agyepong I, Ansah E, et al. Severe Pre-eclampsia adverse Outcome Triage (SPOT) score for of women with pre-eclampsia remote from term: VALidation and development of a risk prediction model (SPOT-VAL study) - study protocol. 2017.

22. Johnston MP. Secondary data analysis: A method of which the time has come. Qual Quant methods Libr. 2017;3(3):619–26.

23. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE):

explanation and elaboration. Int J Surg. 2014;12(12):1500–24.

24. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. Stat Med. 2016;35(2):214–26.

25. Hammond F, Malec J, Nick T, Buschbacher R. Handbook for clinical research: design, statistics, and implementation. Demos Medical Publishing; 2014.

26. White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. Stat Med. 2010;29(28):2920–31.

27. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res. 2011;20(1):40–9.

28. Bennett DA. How can I deal with missing data in my study? Aust N Z J Public Health. 2001;25(5):464–9.

29. Janssens ACJW, Martens FK. Reflection on modern methods: revisiting the area under the ROC curve. Int J Epidemiol. 2020;49(4):1397–403.

30. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143(1):29–36.

31. Bowes D, Hall T, Gray D. Comparing the performance of fault prediction models which report multiple performance measures: recomputing the confusion matrix. In: Proceedings of the 8th international conference on predictive models in software engineering. 2012. p. 109–18.

32. Organization WH, Sciences C for IO of M. International ethical guidelines for health-related research involving humans. Geneva: Council for International Organizations of Medical Sciences; 2016.

33. Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med. 2014;11(10):e1001744.

34. Pavlou M, Ambler G, Seaman S, De Iorio M, Omar RZ. Review and evaluation of penalised regression methods for risk prediction in low-dimensional data with few events. Stat Med. 2016;35(7):1159–77.

35. Cnattingius S, Norman M, Granath F, Petersson G, Stephansson O, Frisell T. Apgar score components at 5 minutes: risks and prediction of neonatal mortality. Paediatr

Perinat Epidemiol. 2017;31(4):328–37.

36. Lee HC, Subeh M, Gould JB. Low Apgar score and mortality in extremely preterm neonates born in the United States. Acta Paediatr. 2010;99(12):1785–9.

37. Ukah UV, Payne B, Lee T, Magee LA, von Dadelszen P. External validation of the fullPIERS model for predicting adverse maternal outcomes in pregnancy hypertension in low-and middle-income countries. Hypertension. 2017;69(4):705–11.

38. The American Academy of Pediatrics. The Apgar Score. Pediatrics [Internet]. 2015 Oct 1;136(4):819 LP – 822. Available from: http://pediatrics.aappublications.org/content/136/4/819.abstract

39. Doolan DM, Froelicher ES. Using an existing data set to answer new research questions: a methodological review. Res Theory Nurs Pract. 2009;23(3):203–15.

40. Power RA, Muthén B, Henigsberg N, Mors O, Placentino A, Mendlewicz J, et al. Non-random dropout and the relative efficacy of escitalopram and nortriptyline in treating major depressive disorder. J Psychiatr Res [Internet]. 2012;46(10):1333–8. Available from: https://www.sciencedirect.com/science/article/pii/S0022395612001963

## Tables and figures

Table 1:

*Baseline characteristics of newborns and women*

| Characteristics (N= 565) | Total n (%) / mean (SD[d]) | Total missing n (%) | AS[a] < 7 n (%) / mean (SD[d]) N=150 | AS[a] ≥ 7 n (%) / mean (SD[d]) N=341 | P-value* |
|---|---|---|---|---|---|
| | | *Newborns* | | | |
| Sex | | 159 (28.1) | | | .122 |
| Female | 214 (37.9) | | 82 (54.7) | 120 (35.2) | |
| Male | 192 (34.0) | | 60 (40.0) | 122 (35.8) | |
| GA[b] at delivery (weeks) | (33.3 ± 3.4) | 35 (6.2) | (31.0 ± 2.6) | (34.5 ± 3.1) | **.000** |
| Birth weight (grams) | (1840 ± 832) | 71 (12.6) | (1265 ± 448) | (2098 ± 838) | **.000** |
| Congenital anomalies | 5 (0.9) | 114 (20.2) | 3 (2.0) | 2 (0.6) | .163 |
| Resuscitation/ventilation >1min | 89 (15.8) | 89 (15.8) | 51 (34.0) | 36 (10.6) | **.000** |
| Admission NICU[c] | 398 (70.4) | 51(9.0) | 146 (97.3) | 231 (67.7) | **.000** |
| Survival before hospital discharge | 393 (69.6) | 145(25.7) | 134 (89.3) | 231 (67.7) | .852 |
| Mortality within 6 weeks | 96 (17.0) | 168(29.7) | 54 (36) | 35 (10.3) | **.000** |
| Serious morbidities within 6 weeks | 28 (5.0) | 198(35.0) | 17 (11.3) | 9 (2.6) | **.002** |
| | | *Women* | | | |
| Maternal age (years) | (32.2 ± 5.7) | 20 (3.5) | (31.8 ± 5.6) | (32.2 ± 5.7) | .509 |
| Ethnic origin | | 20 (3.5) | | | .089 |
| Akan | 287 (50.8) | | 84 (56.0) | 166 (48.7) | |
| Ewe | 82 (14.5) | | 16 (10.7) | 51 (15) | |
| Ga-Dagma | 105 (18.6) | | 22 (14.7) | 70 (20.5) | |
| Northern | 43 (7.6) | | 17 (11.3) | 22 (6.5) | |
| Hausa | 24 (4.2) | | 4 (2.7) | 17 (5.0) | |
| Other | 4 (0.7) | | 2 (1.3) | 2 (0.6) | |
| Religion | | 13 (2.3) | | | .750 |
| Christian | 497 (88.0) | | 129 (86.0) | 300 (88.0) | |
| Muslim | 53 (9.4) | | 16 (10.7) | 33 (9.7) | |
| Traditional | 2 (0.4) | | - | 1 (0.3) | |
| Other | - | | - | - | |
| Educational status | | 16 (2.8) | | | .245 |
| No formal education | 30 (5.3) | | 10 (6.7) | 13 (3.8) | |
| Primary | 51 (9.0) | | 17 (11.3) | 23 (6.7) | |
| JSS | 232 (41.1) | | 61 (40.7) | 146 (42.8) | |
| SSS | 112 (19.8) | | 25 (16.7) | 75 (22.0) | |
| Tertiary | 121 (21.4) | | 28 (18.7) | 76 (22.3) | |
| Other | 3 (0.5) | | 1 (0.7) | 2 (0.6) | |
| Marital status | | 19 (3.4) | | | .159 |
| Married | 432 (76.5) | | 104 (69.3) | 269 (78.9) | |
| Single | 102 (18.1) | | 35 (23.3) | 57 (16.7) | |
| Cohabiting | 12 (2.1) | | 3 (2.0) | 6 (1.8) | |
| Widowed | - | | - | - | |
| Other | - | | - | - | |
| Nicotine use (during pregnancy) | 1 (0.2) | 13 (2.3) | 1 (0.7) | - | .128 |
| Alcohol use (during pregnancy) | 20 (3.5) | 43 (7.6) | 6 (4.0) | 12 (3.5) | .777 |
| Drug use (during pregnancy) | 1 (0.2) | 27 (4.8) | - | - | - |

Percentages may not sum up to 100% because of rounding and missing values within the Apgar score. The Apgar score has 74 missing values.
* : A *P* value <0.05 is considered to be statistically significant.
a: Apgar score, b: Gestational age, c: Neonatal intensive care unit, d: Standard deviation

Table 2:

*Missing values*

| Characteristics (N= 565) | Total missing n (%) | Missing AS < 7 n (%) | Missing AS ≥7 n (%) |
|---|---|---|---|
| | *Newborns* | | |
| Sex | 159(28.1) | 8(5.3) | 99(29.0) |
| GA at delivery (weeks) | 35(6.2) | - | 3(0.9) |
| Birth weight (grams) | 71(12.6) | 2(1.3) | 10(2.9) |
| Congenital anomalies | 114(20.2) | 19(12.7) | 53(15.5) |
| Resuscitation/ventilation >1min | 89(15.8) | 10(6.7) | 26(7.6) |
| Admission NICU | 51(9.0) | 1(0.7) | 5(1.5) |
| Survival before hospital discharge | 145(25.7) | 8(5.3) | 95(27.9) |
| Mortality within 6 weeks | 168(29.7) | 14(9.3) | 115(33.7) |
| Serious morbidities within 6 weeks | 198(35.0) | 25(16.7) | 131(38.4) |
| | *Women* | | |
| Maternal age (years) | 20(3.5) | 5(3.3) | 13(3.8) |
| Ethnic origin | 20(3.5) | 5(3.3) | 13(3.8) |
| Religion | 13(2.3) | 5(3.3) | 7(2.1) |
| Educational status | 16(2.8) | 8(5.3) | 6(1.8) |
| Marital status | 19(3.4) | 8(5.3) | 9(2.6) |
| Nicotine use (during pregnancy) | 13(2.3) | 5(3.2) | 6(1.8) |
| Alcohol use (during pregnancy) | 43(7.6) | 12(8.0) | 24(7.0) |
| Drug use (during pregnancy) | 27(4.8) | 6(4.0) | 17(5.0) |

Percentages may not sum up to 100% because of rounding and missing values within the Apgar score. The Apgar score has 74 missing values.

Table 3:

*Univariate & multivariate logistic regression analysis*

| Variable | Univariate | | Multivariate | |
|---|---|---|---|---|
| | P-value | OR (95% CI) | P-value | AOR (95% CI) |
| | | *NICU admission* | | |
| APGAR (0 - 10) | *** | 2.23 (1.79-2.87) | - | |
| Sex | *** | 0.40 (0.26-0.62) | - | |
| GA at delivery (weeks) | *** | 1.93 (1.73-2.19)[a] | - | |
| Birth weight (grams) | *** | 1.002 (1.0019-1.0028)[a] | - | |
| Maternal age (years) | *** | 1.07 (1.03-1.01) | * | 1.09(1.02-1.16) |
| Ethnic origin | - | - | - | |
| Religion | - | - | - | |
| Educational status | - | - | - | |
| Marital status | - | - | - | |
| Alcohol use (during pregnancy) | - | - | | |
| | | *Resuscitation/ventilation* | | |
| APGAR (0 - 10) | *** | 1.55(1.36-1.77) | *** | 1.42(1.23-1.67) |
| Sex | - | - | | |
| GA at delivery (weeks) | *** | 1.202(1.115-1.301) | | |
| Birth weight (grams) | *** | 1.001(1.0006-1.0013) | | |
| Maternal age (years) | ** | 1.06(1.02-1.1) | ** | 1.06(1.02-1.11) |
| Ethnic origin | - | - | | |
| Religion | - | - | | |
| Educational status | - | - | | |
| Marital status | - | - | | |
| Alcohol use (during pregnancy) | - | - | | |
| | | *Morbidities within 6 weeks* | | |
| APGAR (0 - 10) | ** | 1.32(1.10-1.59) | * | 1.24(1.00-1.52) |
| Sex | - | - | | |
| GA at delivery (weeks) | ** | 1.25(1.09-1.45) | * | 1.267(1.02-1.57) |
| Birth weight (grams) | - | - | | |
| Maternal age (years) | - | - | | |
| Ethnic origin | - | - | | |
| Religion | - | - | | |
| Educational status | - | - | | |
| Marital status | - | - | | |
| Alcohol use (during pregnancy) | - | - | | |
| | | *Mortality within 6 weeks* | | |
| APGAR (0 - 10) | *** | 1.49(1.30-1.71) | *** | 1.31(1.23-1.53) |
| Sex | - | - | | |
| GA at delivery (weeks) | *** | 1.36(1.24-1.50)[a] | *** | 1.52(1.24-1.90) |
| Birth weight (grams) | *** | 1.002(1.0012-1.0024)[a] | *** | 1.011 (1.006-1.016) |
| Maternal age (years) | - | - | | |
| Ethnic origin | - | - | | |
| Religion | - | - | | |
| Educational status | - | - | | |
| Marital status | - | - | | |
| Alcohol use (during pregnancy) | - | - | | |

Significant codes: 0 '***'; 0.001 '**' ;0.01 '*' ; > 0.05 '-'

a: An interaction term between GA and birth weight is added in the multivariate analyses.

GA and the significant variables from the univariate analysis were selected as a confounder(s) and entered in the multivariate model

Table 4:

*Prediction of adverse outcomes*

| Adverse outcome | SENS | SPEC | NPV | PPV | AUC univariate[a] (95% CI) | AUC multivariate[b] (95% CI) |
|---|---|---|---|---|---|---|
| Resuscitation/ventilation >1min | 0.43 | 0.23 | 0.64 | 0.11 | 0.73(0.68-0.79) | 0.76(0.71-0.82) |
| Admission NICU | 0.63 | 0.03 | 0.02 | 0.69 | 0.78(0.73-0.82) | - |
| Morbidity within 6 weeks | 0.39 | 0.32 | 0.86 | 0.05 | 0.67(0.56-0.78) | 0.72(0.62-0.81) |
| Mortality within 6 weeks | 0.44 | 0.27 | 0.60 | 0.16 | 0.70(0.63-0.76) | 0.78(0.72-0.84) |

a: AUC scores between Apgar and an adverse outcome in a univariate analysis
b: AUC scores between the complete multivariate model and the adverse outcome

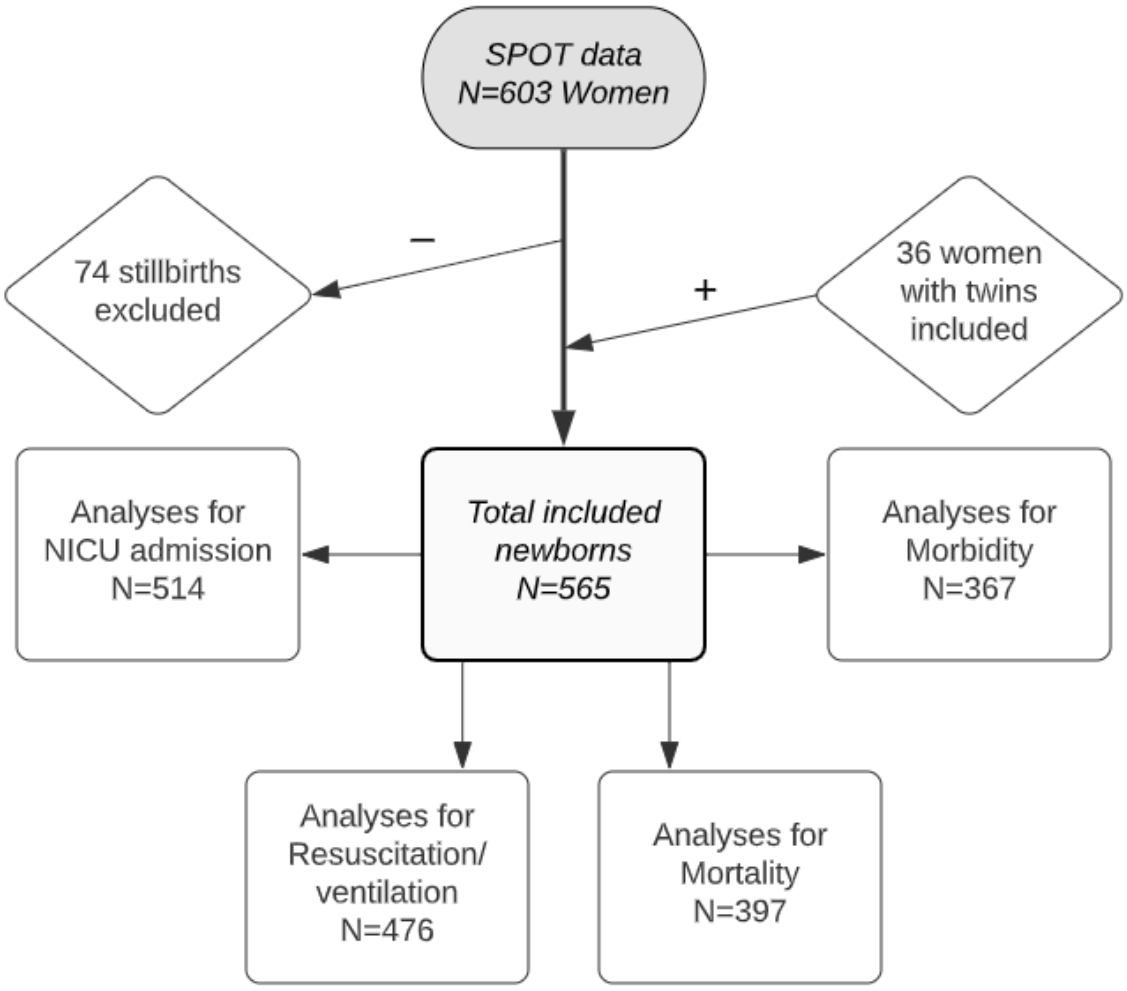


*Figure 1:* Flowchart

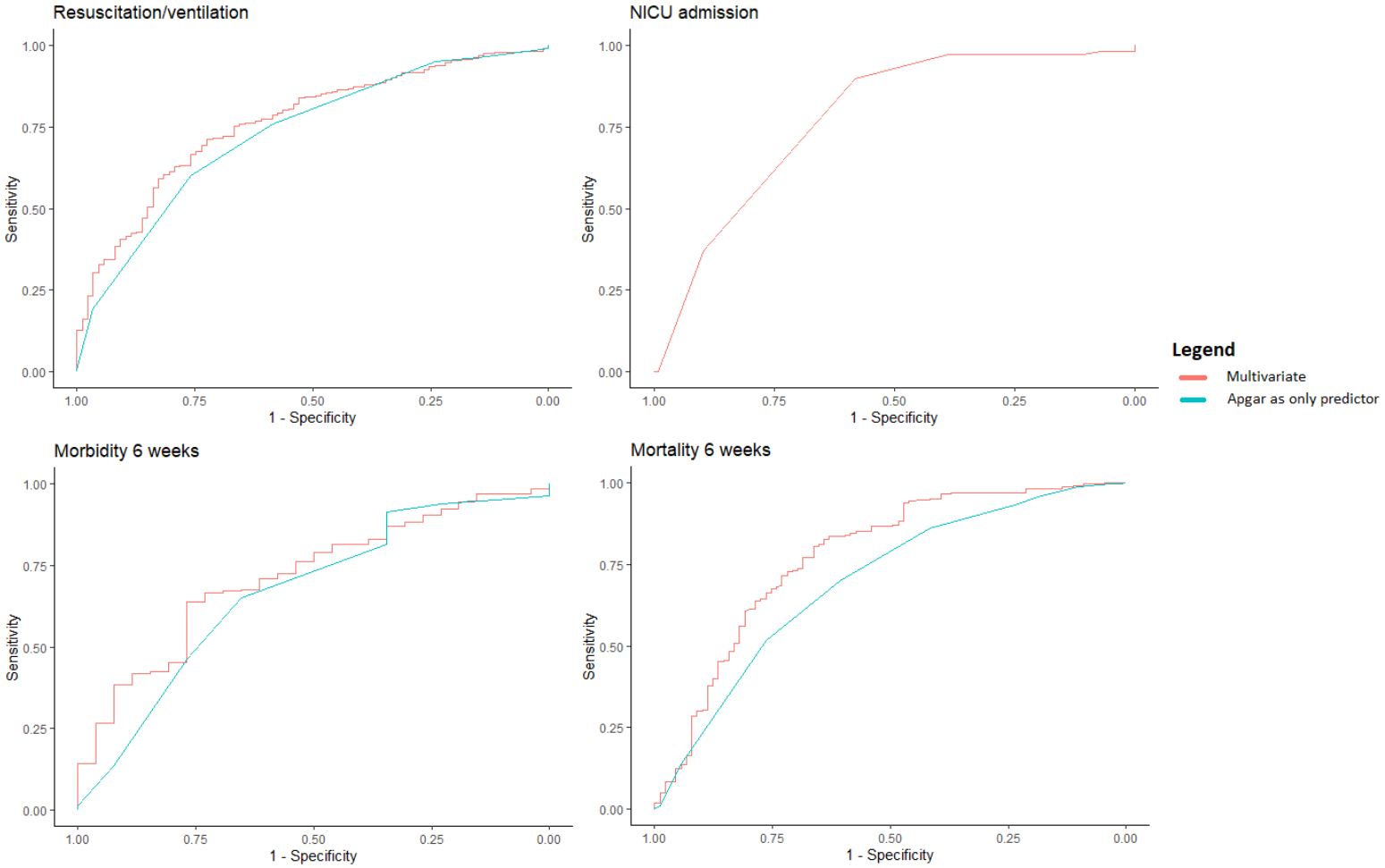


*Figure 2*: ROC-curves