scientific reports

OPEN

Neonatologist staffing is related to the inter-hospital variation of risk-adjusted mortality of very low birth weight infants in Korea

Myung Hee Lee[1,5], Jang Hoon Lee[2,5] & Yun Sil Chang[3,4]✉

This study investigated whether hospital factors, including patient volume, unit level, and neonatologist staffing, were associated with variations in standardized mortality ratios (SMR) adjusted for patient factors in very-low-birth-weight infants (VLBWIs). A total of 15,766 VLBWIs born in 63 hospitals between 2013 and 2020 were analyzed using data from the Korean Neonatal Network cohort. SMRs were evaluated after adjusting for patient factors. High and low SMR groups were defined as hospitals outside the 95% confidence limits on the SMR funnel plot. The mortality rate of VLBWIs was 12.7%. The average case-mix SMR was 1.1; calculated by adjusting for six significant patient factors: antenatal steroid, gestational age, birth weight, sex, 5-min Apgar score, and congenital anomalies. Hospital factors of the low SMR group (N = 10) had higher unit levels, more annual volumes of VLBWIs, more number of neonatologists, and fewer neonatal intensive care beds per neonatologist than the high SMR group (N = 13). Multi-level risk adjustment revealed that only the number of neonatologists showed a significant fixed-effect on mortality besides fixed patient risk effect and a random hospital effect. Adjusting for the number of neonatologists decreased the variance partition coefficient and random-effects variance between hospitals by 11.36%. The number of neonatologists was independently associated with center-to-center differences in VLBWI mortality in Korea after adjustment for patient risks and hospital factors.

Keywords Infant, Very low birth weight, Mortality, Premature, Standardized mortality ratio

Although advances in neonatal intensive care have markedly improved the outcomes of very-low-birth-weight infants (VLBWIs) weighing less than 1500 g at birth[1,2] mortality rates still vary among neonatal intensive care units (NICUs), even in many developed countries[3–7]. The mortality rates of VLBWIs are regarded as objective markers of the quality of care provided in NICUs[8]. However, differences in patient and hospital factors can complicate the direct comparisons of outcomes between NICUs[9,10]. Furthermore, concerns exist that comparative mortality measures in acute medical settings can result in unfair judgments and institutional stigmatization, diverting focus from genuine to superficial improvements. Therefore, the over or misinterpretation of these results should be avoided[11].

To solve this issue, a case-mix standardized mortality ratio (SMR) has been developed to adjust for these high-risk infant characteristics[12], including known essential risk factors such as maternal factors, perinatal factors, congenital malformations, and low birth weight[13–17]. SMR allows for the mortality rates of each center to be compared to those of other centers[7,18–22]. Furthermore, multi-level risk adjustments with both patient and hospital factors can reliably provide both a more precise, unbiased assessment of outcomes and an acceptable center-to-center variation that may be used as a benchmark indicator for quality improvement[10,23].

However, the data on the impact of hospital factors on preterm infant mortality are controversial. The annual volume of VLBWIs has been shown to both affect[24–28] and not affect[8,9,29] mortality outcomes, whereas the unit level usually affected mortality[8,26,28]. In contrast, neonatologist coverage[30,31] or residency programs[9,31] did not

[1]Institute of Biomedical and Clinical Research, MEDITOS, Seoul, Republic of Korea. [2]Department of Pediatrics, Ajou University School of Medicine, Suwon, Korea. [3]Department of Pediatrics, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-Ro, Gangnam-Ku, Seoul 06351, Korea. [4]Department of Health Sciences and Technology, Samsung Advanced Institute for Health Sciences and Technology (SAIHST), Sungkyunkwan University, Seoul, Korea. [5]These authors contributed equally: Myung Hee Lee and Jang Hoon Lee. ✉email: yschang@skku.edu; cys.chang@samsung.com

show a meaningful relationship with VLBWIs outcomes. However, neonatologist staffing below a critical point was closely related to higher mortality among extremely low birth weight infants weighing less than 1,000 g at birth[32]. Furthermore, neonatologist staffing conditions differ across healthcare systems both nationwide and worldwide in terms of how the roles of NICUs are defined and how the medical staff are allocated.

Recently, the perinatal and neonatal medical field in Korea has faced several urgent issues, including an unprecedented ultra-low fertility rate, a relatively increasing rate of preterm births[33], a sharp decline in the number of pediatricians, and difficulties in recruiting and employing perinatal and neonatal specialists due to low reimbursement rates under the single national insurance system[34].

Previously, regional differences in VLBWI mortality in Korea have been shown to be primarily due to differences in neonatal care resources[35]. Moreover, a recent report showed that the number of NICU beds per neonatologist in Korea was 13.4, which is almost twice as high as that reported in the United States (8.4), Australia (6.9), and Canada (8.1)[36] suggests that the supply of neonatologists in Korea falls below international standards[37,38] necessitating further exploration into the effect of neonatologist staffing on NICU outcomes.

The present study aimed to report the multi-level risk-adjusted mortality of VLBWIs with center-to-center variation to promote quality improvement in neonatal intensive care in Korea. To achieve this objective, we first developed a case-mix SMR for VLBWIs using a national prospective cohort from the Korean Neonatal Network (KNN) covering over 80% of the annual VLBWIs in Korea[39]. Second, we compared high and low SMR groups by making funnel plots for case-mix SMR and then investigated which factors, such as patient volume, unit level, and neonatologist staffing, could reduce the center-to-center mortality variation by performing multi-level risk adjustment.

Methods

Data source

We used a de-identified dataset approved by the Committee of Ethics and Publication of the KNN. The dataset was extracted from the internet-based clinical trial management system (i-CReaT) of the Korea National Institute of Health. The KNN registry includes VLBWIs admitted to the NICU at birth or transferred from other hospitals within 28 days of birth, excluding those who died prior to NICU admission. Perinatal and neonatal information was prospectively collected from NICUs using the KNN manual of operation and electronic case report forms based on the i-CReaT system. After completing data quality management using queries and site-visit monitoring[39], the KNN registry data were stored on the Korea National Institute of Health server. Individual hospital information, such as the number of NICU beds, working neonatologists, and various neonatal care resources, was collected annually in the form of standardized survey reports from the principal investigators of each participating hospital.

Study population and analysis sample

We initially included all VLBWIs (N = 16,386) born between January 2013 and December 2020. The exclusion criteria included infants registered at hospitals with < 30 VLBWIs enrollments during the study period (9/72 hospitals, N = 89) and cases with missing data regarding potential patient risk factors for mortality, except for chorioamnionitis (N = 531). A total of 620 cases were excluded and 15,766 cases were analyzed (Fig. S1). Data from VLBWIs born between January 2013 and December 2018 (N = 12,055) were used to develop and internally validate the prediction model. External validation was performed using the dataset of VLBWIs born between January 2019 and December 2020 (N = 3711).

For the subgroup analysis, gestational age was classified as < 26 weeks, 26–28 weeks, 29–31 weeks, and ≥ 32 weeks; birth weight was categorized into four groups: ≤ 749 g, 750–999 g, 1000–1249 g, and 1250–1499 g.

Outcome assessment

The primary outcome was the center-to-center mortality variation among the 63 NICUs, based on the estimated SMR during the hospitalization period up to 365 days after birth. To identify factors affecting mortality variation, risk factors at both the patient and hospital level were assessed using statistical modeling, including a two-level mixed model. The SMR was calculated by dividing a hospital's observed deaths by the expected numbers of deaths. An SMR funnel plot was constructed to identify variations in hospital mortality[40]. SMR beyond the + 95% confidence limit or below -95% confidence limit of the funnel plot were considered to be prospective high or low statistical anomalies, showing evidence of opposing quality of clinical performance[41,42].

The basic characteristics of the study population included perinatal and neonatal variables. Perinatal variables included multiple births, chorioamnionitis, pregnancy-induced hypertension (PIH), gestational diabetes mellitus (GDM), use of antenatal steroids, and mode of delivery (cesarean section/vaginal delivery). Neonatal variables included gestational age (weeks), birth weight (g), sex, small-for-gestational-age, Apgar score (1-min, 5-min), place of delivery (inborn/outborn), and major congenital anomalies. Hospital information included the number of NICU beds, the number of board-certified neonatologists working in the NICU, and various neonatal care resources. In the statistical analysis, we applied data merged by year from the hospital information and KNN registry data. In Korea, neonatologist board certification is available to pediatric specialists who have completed a four-year residency and passed the pediatric specialist exam. They must then complete two years of neonatology training at the Korean Society of Neonatology-designated hospitals, including at least one year as a fellow. After this training, they must pass written and oral exams administered by the Korean Pediatric Society to earn the neonatologist board certification.

To evaluate the level of each NICU, we used a scoring system adapted from the proposed uniform definition of the neonatal intensive care level[43] as previously described[35]. Each of the following nine items was awarded a point: availability of total parenteral nutrition, general pediatric surgery, pediatric thoracic surgery, nitric oxide

therapy, extracorporeal membrane oxygenation therapy, dialysis treatment, echocardiography, other ultrasounds, and blood gas analysis within the unit. Based on their score, each hospital was assigned one of three levels: level 1 for a score of 1–5, level 2 for 6–7, and level 3 for 8–9. Accordingly, hospitals with higher scores had higher levels of neonatal intensive care.

After establishing the case-mix SMR, we divided hospitals into two groups based on their SMR: High SMR group were beyond the +95% confidence limits and Low SMR group were below the -95% confidence limits on the funnel plot.

Statistical analysis

Continuous variables are presented as means ± standard deviation and were compared using the independent t-test or Wilcoxon rank-sum test. Categorical variables are expressed as percentages and frequencies and analyzed using the chi-square or Fisher's exact tests.

We used univariate and multivariate logistic regressions to identify the risk factors associated with death using STATA and the R module "bestglm"[44]. The best model was chosen based on the Bayesian information criterion (BIC), Akaike information criterion (AIC), and C-statistic[45–47]. The model with the lowest BIC and AIC or the highest C-statistic is preferred. Based on the Hosmer–Lemeshow chi-square statistics, the goodness-of-fit test was used to assess calibration[48]. For internal validation, the bootstrapping method to measure the Somers' D rank correlation and the C-statistic for each bootstrapped sample[49,50]. External validation was done with the Hosmer–Lemeshow test, and a calibration plot.

The SMR was calculated as the observed deaths divided by the expected deaths. The adjusted SMR was calculated using a prediction model for each neonate. Funnel plots for the SMR based on the Poisson distribution were used to assess between-hospital variation[51].

To identify the contributors to hospital variation in mortality, we performed a multi-level mixed-effect logistic regression, considering random hospital effect, patient risk effect, and hospital-level covariates[52,53]. The variation measures included the proportional change in the between-hospital variance (PCV), variance partition coefficient (VPC), and median odds ratio (MOR). These variables were calculated as follows: PCV = $\{(V_A - V_B)/V_A\} \times 100$, where V_A is the variance of initial model and V_B is the variance of the model with more covariates, such as patient- and hospital-level variables; VPC = $V_C/(V_C + 3.29) \times 100$, where V_C is the residual variance of the hospital-level; and MOR = $\exp\left(0.95 \times \sqrt{V_C}\right)$[54,55].

A higher VPC and MOR indicate greater between-hospital variation. A higher PCV means the added covariates reduced inter-hospital variation.

Statistical analyses were performed using STATA 14.0 and R version 4.1.1. P-values < 0.05 were considered statistically significant for all analyses.

Ethical approval

The Institutional Review Board approved the KNN registry at each participating hospital, and informed consent was obtained from the parents upon admission to the participating NICUs. All methods were performed in accordance with relevant guidelines and regulations.

Results

Patient demographics

Patient demographics regarding perinatal and neonatal factors and mortality according to gestational age and birth weight are shown in Table S1. Among the 15,766 patients, the overall unadjusted hospital mortality rate was 12.7% (N = 1,997); 42.8% of deaths (N = 856) occurred within 7 days of birth, 59.3% (N = 1185) within 14 days, and 76.6% (N = 1,530) within 28 days.

The mortality rates according to birth weight were 42.2% (N = 1031) for < 750 g, 16.7% (N = 582) for 750–999 g, 5.6% (N = 239) for 1,000–1,249 g, and 2.5% for 1,250–1,499 g. The mortality rate according to gestational age was 41.4% (N = 1,138) for < 26 weeks, 11.6% (N = 604) for 26–28 weeks, 3.7% (N = 200) for 29–31 weeks, and 2.1% (N = 55) for > 31 weeks.

Model development and external validation for SMR

Table S2 presents a comparison of the characteristics between survivors and non-survivors, along with the logistic regression for predicting in-hospital mortality in a model development dataset comprising 12,055 patients. Univariate logistic regression analysis identified several predictors of mortality, including chorioamnionitis, PIH, GDM, antenatal steroid use, cesarean section, gestational age, birth weight, Apgar score (1-min, 5-min), and major congenital anomalies.

Table S3 displays the results of the multiple prediction models (models 1–6) developed by multiple logistic regression to select the best predictive model using 13 candidate predictors, excluding chorioamnionitis owing to missing values. The final prediction model, model 5, included the following variables: antenatal steroid use, gestational age, birth weight, sex, 5-min Apgar score, and major congenital anomalies. Model 5 was chosen based on its goodness-of-fit (Hosmer–Lemeshow test P-value = 0.6838), robust prediction performance with a C-statistic of 0.86 (95% confidence interval (CI), 0.85–0.87), similar performance to models 1–4 despite fewer predictors, and BIC value being lowest among the six models.

The internal validation of the prediction model using bootstrapping showed that the estimated Somers' D correlation coefficients and C-statistics of the bootstrap models were comparable to those of the development model, and the difference between the estimated parameters of the two models was approximately 0. External validation using an independent dataset confirmed the fitness of the risk prediction model (Fig. S2).

Table 1 presents the characteristics of survivors and non-survivors as well as the logistic regression for all VLBWIs (N = 15,766) based on the predictors selected in the final prediction model, which showed little difference compared with the model development sets (N = 12,055, Table S2).

Comparisons between hospitals with high and low SMR

The average risk-adjusted SMR of the KNN hospitals over an 8-year period was 1.1 ± 0.46 (range, 0.25–2.21). The distribution of SMR was as follows: < 0.75 (15 NICUs, 23.8%); 0.75–1.25 (26 NICUs, 41.3%); 1.25–1.50 (9 NICUs, 14.3%); 1.50–2.0 (11 NICUs, 17.5%); and > 2.0 (2 NICUs, 3.2%). Analysis of case-mix SMR adjusted for gestational age and birth weight revealed significant variations among the 63 hospitals (Fig. S3). For all VLBWIs, 13 hospitals (High SMR) had SMR values beyond the + 95% confidence limit, whereas 10 hospitals (Low SMR) had values below the -95% confidence limit. The subgroup funnel plot showed a tendency for increased SMR variation as the gestational age or birth weight decreased. Table S4 provides a description and comparison of the characteristics of patients and hospital factors between infants born in the High SMR hospitals (Group B) and infants born in Low SMR hospitals (group A) based on the funnel plot (Fig. S3A). There were no differences in mean gestational age and birth weight between the two groups, and Group B exhibited significantly lower mortality rates than did Group A. The incidences of multiple births, chorioamnionitis, and antenatal steroid use were higher in Group B. Furthermore, Group B had lower Apgar scores at 1 and 5 min, a lower rate of infants born in hospitals, and a higher rate of congenital anomalies. Notably, significant differences were observed between the two groups in terms of unit-level factors. Group B had a higher level of care and higher numbers of VLBWIs admitted per year, neonatologists, and VLBWIs per neonatologist, as well as a lower number of NICU beds per neonatologist. Fig. S4 displays the OR for mortality and major morbidities between Groups A and B. Group A had a significantly higher risk of mortality (OR 8.12, 95% CI 6.64–9.93, P-value < 0.001) and mortality or morbidities (OR 2.56, 95% CI 2.26–2.90, P-value < 0.001) than Group B.

Hospital-level factors influencing inter-hospital variation of SMR

The funnel plot for the case-mix adjusted SMR incorporating both patient- and hospital-level factors reveals a reduction in inter-hospital variation, particularly beyond the + 95% Confidence limit, from 13 hospitals in the pre-adjusted model to 9 hospitals after adjustment (Fig. 1).

To further investigate the inter-hospital variation in mortality in VLBWIs, we utilized a multi-level model. The multi-level mixed-effect logistic regression analysis that considered hospital factors, including the level of care, number of VLBWIs, and number of neonatologists, in addition to patient factors, is presented in Table 2.

Significantly higher mortality risk was observed when considering patient-level risk factors in the entire cohort (OR 3.0, 95% CI 2.87–3.15, P-value < 0.001). Younger gestational age and lower birth weight were associated with an increased mortality risk.

Since including Apgar scores in patient-level risk adjustment could inadvertently favor hospitals with suboptimal resuscitation quality by artificially lowering their risk-adjusted SMRs, we conducted a sensitivity analysis excluding them. The results (Table S5) were consistent with the original analysis, which included the 5-min Apgar scores (Table 2). Level of care and number of VLBWIs admitted per year did not influence VLBWIs mortality. However, the number of neonatologists had a significant effect on VLBWIs mortality. As gestational age and birth weight decreased, the VPC and MOR showed higher numbers, indicating larger variations (Table 2) and increased inter-hospital variation (Table S6).

Adjusting for the number of neonatologists resulted in a significant reduction in inter-hospital variation, as evidenced by an increased PCV of 11.36%. This indicates an 11.36% reduction in inter-hospital variability compared with the reference model, which was only adjusted for patient risks. The magnitude of this reduction in variation was consistent with the PCV of 11.36% observed in the model adjusted for all three hospital factors. In the subgroup analysis, this reduction of inter-hospital variation was statistically significant for infants born with birth weights below 1,250 g and gestational ages not exceeding 28 weeks. These findings suggest that the number of neonatologists played a crucial role in explaining the inter-hospital variation in mortality of VLBWIs.

Figure 2 depicts the VPC between hospitals stratified by gestational age and birth weight. Overall, the VPC increased after adjusting for patient-related risk factors but decreased after further adjustment for the number of neonatologists.

Variable	Survivor (n = 13,769)	Non-survivor (n = 1997)	P-value	Adjusted[a] OR (95% CI)	P-value
Perinatal & Neonatal					
Antenatal steroids	11,523(83.7)	1562(78.2)	< 0.001	0.57(0.50–0.66)	< 0.001
Gestational age	$29^{+2} \pm 2^{+6}$	$26^{+1} \pm 2^{+3}$	< 0.001	0.78(0.75–0.81)	< 0.001
Birth weight (g)	1125.3 ± 260.3	786.6 ± 254.8	< 0.001	0.99(0.99–0.99)	< 0.001
Male sex	6792(49.3)	1088(54.5)	< 0.001	1.34(1.20–1.50)	< 0.001
5- min Apgar score	7.1 ± 1.6	5.4 ± 2.1	< 0.001	0.82(0.80–0.84)	< 0.001
Congenital anomalies	374(2.7)	178(8.9)	< 0.001	6.99(5.51–8.85)	< 0.001

Table 1. Characteristics of survivors and non-survivors and multivariate logistic regression for death in all VLBWIs (N = 15,766). CI, Confidence interval; OR, Odds ratio; VLBWIs, Very low birth weight infants. Values are presented as means ± SD or n (%). [a]Logistic regression model for the mortality of all VLBWIs based on the selected risk factors.
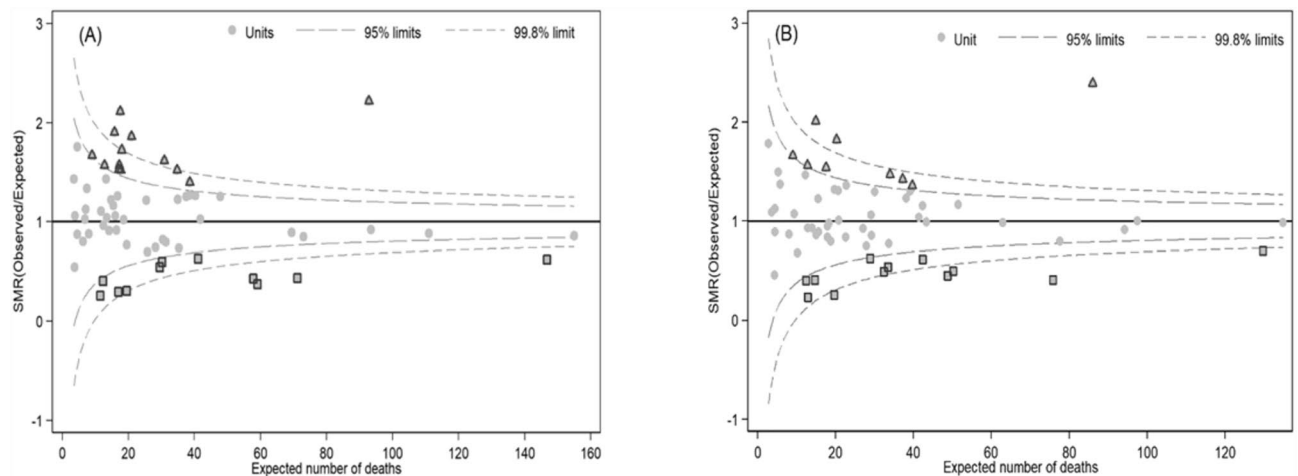
Figure 1. Case-mix adjusted standardized mortality ratio (SMR) for all VLBWIs. These funnel plots presented adjusted SMR between patient level and hospital level. A was adjusted by only risk factors of patient level (gestational age, birth weight, gender, 5-min Apgar score, antenatal steroid and congenital anomalies). B was adjusted by risk factor of patient level and all hospital level (level of care, number of VLBWIs per year and number of neonatologists). The triangles represent that adjusted SMR is higher than expected mortality at 95% confidence limit, and the squares represent that adjusted SMR is lower than expected mortality at 95% confidence limit.

When plotting the patient-risk-adjusted SMR against the number of neonatologists, a clear trend of decreasing SMR was observed as the number of neonatologists increased. Importantly, the trend slopes between the survival and death groups were significantly different, indicating a closer relationship in the death group. The SMR approached 1 when the death group had ≥ 4 neonatologists, and the survival group showed a similar trend with ≥ 2 neonatologists (Fig. 3).

Discussion

In the present study, we have shown considerable inter-hospital variation in the mortality rates of VLBWIs by obtaining SMRs adjusted for six important patient risk factors from 63 NICUs in Korea. Furthermore, we found that the number of neonatologists working in the NICU had a more significant impact on this variation than the unit level of care or patient volume admitted within a given year.

VLBWIs mortality rates in NICUs exhibit regional and hospital variations[7,9,12], making it crucial to address these outcome differences as part of quality improvement initiatives[41]. SMR adjusted for patient characteristics has been used as a benchmark for comparing outcomes among hospitals because hospital-admitted patient populations cannot be randomly controlled[9,12,18]. In the present study, we calculated adjusted SMRs to compare NICU quality indicators across hospitals in Korea and showed inter-hospital variations using funnel plots.

In an Australian network study, a significant reduction in inter-hospital variation of observed mortality was shown when adjusting only for patient characteristics, with all hospitals falling within the ± 99.8 percentile range on the funnel plot[20]. However, our study revealed considerable variability in the funnel plot of SMR adjusted for patient characteristics, with more prominent variations observed as the gestational age and birth weight decreased. This can be attributed to differences in the study settings, as the Australian study focused on eight relatively homogeneous tertiary units in a similar region, whereas our study included 63 nationwide hospitals with varying levels of care. This phenomenon has been consistently observed not only in other studies focusing on mortality as an outcome[21,56–59] but also in studies examining other outcomes within the NICU, such as severe intraventricular hemorrhage[10], nosocomial infection[29], and bronchopulmonary dysplasia[60]. Therefore, controlling for patient characteristics alone does not resolve inter-center variations, highlighting the need for further adjustment in the next stage of analysis. Furthermore, our study demonstrated that when adjusting for patient characteristics, inter-hospital variation increased, which aligns with similar patterns observed in other studies on intraventricular hemorrhage[10]. This suggests the importance of case-mix adjustment in comparing centers and the necessity of further analysis of other hospital factors.

Moreover, caution is required when comparing outcomes based solely on adjusted SMR and funnel plots, as patients are nested within hospitals, necessitating the consideration of hospital factors in the statistical analysis[23]. Hospitals with a small number of cases are not suitable for SMR comparisons, and adjustments should be made to account for both random effects and actual differences between hospitals[29,59]. Thus, we excluded hospitals with fewer than 30 enrolled cases during the study period and conducted analyses using multi-level modeling to obtain adjusted SMRs that accounted for hospital factors.

Identifying suitable hospital factors that can effectively explain inter-hospital variations in NICU care remains challenging. Possible factors include the level of care; staffing of neonatologists, nurses, and other health professionals; the practice quality of each unit; quality improvement initiatives; networking through referral systems; and parental involvement. However, appropriate measurement of these factors is difficult and complex. Herein,

Variable	Birthweight (g)				Gestational age (weeks)				Total
	≤749	750–999	1000–1249	1250–1499	≤25	26–28	29–31	≥32	
	(N=2441)	(N=3481)	(N=4230)	(N=5614)	(N=2748)	(N=5204)	(N=5282)	(N=2532)	(N=15,766)
Fixed effects[a]									
Risk score[b]	3.19 (2.79–3.64)	3.15 (2.76–3.60)	2.99 (2.60–3.43)	2.81 (2.48–3.19)	3.24 (2.84–3.71)	3.15 (2.81–3.52)	2.88 (2.55–3.25)	2.81 (2.32–3.40)	3.00 (2.87–3.15)
Hospital									
Level[c]									
1	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
2	0.85 (0.45–1.58)	1.13 (0.66–1.92)	1.39 (0.71–2.73)	0.77 (1.64–0.36)	0.99 (0.56–1.75)	0.99 (0.61–1.61)	2.06 (0.92–4.61)	0.72(0.20–2.63)	0.97(0.69–1.36)
3	0.88(0.44–1.77)	1.77(0.98–3.20)	1.95(0.92–4.11)	0.97(0.42–2.22)	1.44(0.76–2.74)	1.35(0.78–2.32)	2.05(0.86–4.92)	1.09(0.28–4.24)	1.30(0.88–1.90)
No. of VLBW infants (per year)									
< 20	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
20–40	1.20(0.79–1.80)	1.11(0.75–1.63)	0.72(0.44–1.16)	0.86(0.47–1.56)	1.23(0.84–1.79)	1.04(0.73–1.49)	0.62(0.37–1.04)	0.59(0.21–1.68)	0.98(0.78–1.24)
40–60	1.04(0.63–1.71)	1.67(1.02–2.73)	0.95(0.50–1.79)	0.42(0.18–0.96)	1.36(0.86–2.17)	1.28(0.81–2.03)	0.54(0.27–1.08)	0.22(0.05–1.01)	1.16(0.86–1.56)
≥60	0.81(0.47–1.40)	1.17(0.68–1.99)	1.28(0.67–2.46)	1.07(0.51–2.24)	0.81(0.48–1.36)	1.47(0.90–2.40)	0.66(0.34–1.27)	1.10(0.35–3.48)	0.99(0.70–1.39)
No. of neonatologists									
≤1	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
2	0.87(0.60–1.26)	0.58(0.41–0.83)	0.38(0.24–0.60)	0.96(0.55–1.69)	0.64(0.45–0.90)	0.62(0.45–0.86)	0.79(0.48–1.30)	0.84(0.34–2.07)	0.69(0.55–0.87)
≥3	0.57(0.34–0.95)	0.26(0.16–0.42)	0.29(0.15–0.54)	0.83(0.39–1.76)	0.35(0.22–0.57)	0.32(0.20–0.50)	0.81(0.42–1.57)	0.26(0.08–0.88)	0.42(0.31–0.57)
Random effects									
PCV (%)									
Plus LVNP[d]	− 20.37	− 31.58	− 14.81	− 18.18	− 19.35	− 24.14	3.70	–	− 50.00
Plus PL[d]	− 29.63	− 39.47	− 74.07	− 54.55	− 30.65	− 55.17	− 25.93	–	− 73.08
Plus PV[d]	− 24.07	− 34.21	− 66.67	− 18.18	− 17.74	− 55.17	− 14.81	–	− 65.38
Plus PN[d]	− 16.67	− 15.79	− 29.63	− 40.91	− 16.13	− 27.59	− 29.63	–	− 50.00
Plus P[d]	− 27.78	− 28.95	− 70.37	− 45.45	− 25.81	− 48.28	− 37.04	–	− 69.23
Null model[d]	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref[e]	Ref
Plus LVNP[d]	5.80	− 2.04‡	32.61†	18.75	5.13‡	16.28‡	29.73	–	11.36‡
Plus PL[d]	− 1.45	− 8.16	− 2.17	− 6.25	− 3.85	− 4.65	8.11	–	− 2.27
Plus PV[d]	2.90	− 4.08	2.17	18.75*	6.41	− 4.65	16.22	–	2.27
Plus PN[d]	8.70*	10.20‡	23.91‡	3.13	7.69‡	13.95‡	5.41	–	11.36‡
Plus P[d]	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref[e]	Ref
VPC(%)	16.5	13.2	8.6	7.3	18.4	9.9	7.3	1.5	10.6
MOR	2.15	1.96	1.70	1.62	2.26	1.77	1.62	1.24	1.81

Table 2. Multi-level mixed-effects logistic regression for VLBWIs mortality according to gestational age and birth weight. ICC, intraclass correlation; MOR, median odds ratio; PCV, proportional change in variance; Ref, reference group; SE, standard error; VLBWIs, very low birth weight infants; VPC, variance partition coefficient. Significance of model comparison (chi-square statistic) for effects of adding hospital-level covariates to the null model and to the Plus P model at the patient level: *P-value < .05; †P-value < .01; ‡P-value < .001. [a]These values are presented as odds ratios (ORs) and 95% confidence intervals (CIs). [b]The predicted probability based on risk factors at the patient level (gestational age, birth weight, sex, 5-min Apgar score, antenatal steroid use, and congenital anomalies). [c]Each of the following nine items was given a point: availability of total parenteral nutrition, general pediatric surgery, pediatric thoracic surgery, nitric oxide therapy, extracorporeal membrane oxygenation therapy, dialysis treatment, echocardiography, other ultrasounds, and capability of blood gas analysis within the unit. Each hospital was assigned one level based on the following classification scheme: a score of 1–5 as level 1; a score of 6–7 as level 2; and a score of 8–9 as level 3. [d]Null model, random intercept only model without covariates; Plus P, adjusted for risk factors at the patient level; Plus PN, adjusted for patient risk factors and the number of neonatologists (personnel); Plus PV, adjusted for patient risk factors and the number of VLBW infants per year (volume); Plus PL, adjusted for patient risk factors and facility level (level); Plus LVNP, adjusted for patient risk factors, level, volume, and personnel. The significance of model comparison for adding hospital-level and patient-level covariates to the null model is P-value < .001 in all models. [e]For infants with a gestational age of ≥ 32 years, we did not calculate PCV with respect to the null model or the Plus P model because the between-hospital random effect is close to zero (P-value ≥ .05).

we compared hospitals outside the ± 95% CI on the case-mix-adjusted SMR funnel plot and identified three factors with the most significant differences: level of care, patient volume, and number of neonatologist staff.
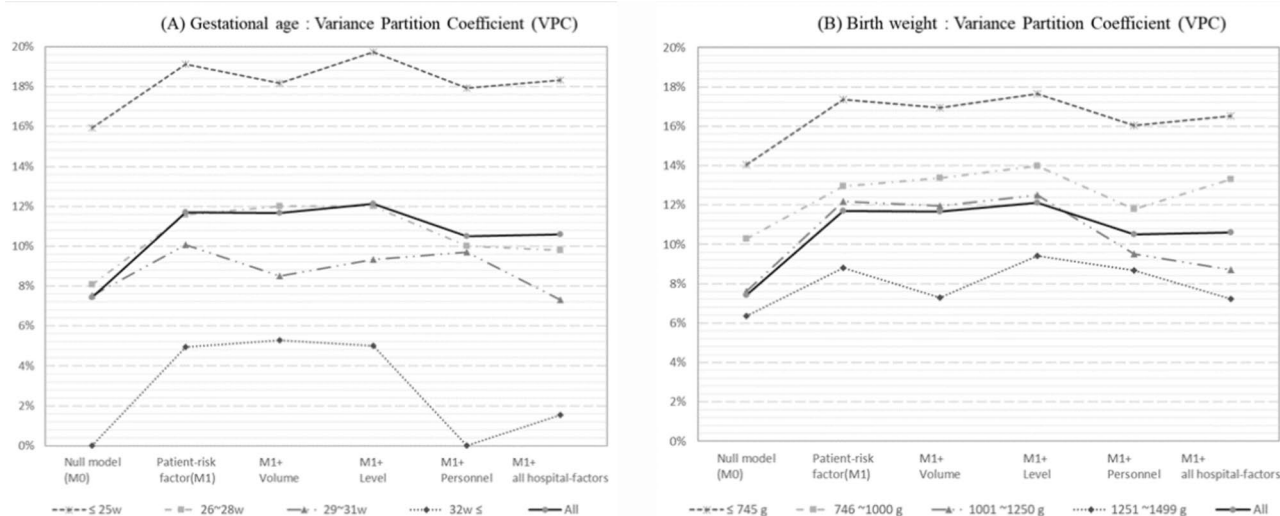
Figure 2. Variance partition coefficient between hospitals by gestational age and birth weight. All risk factors of hospital level presented "Level + Volume + Personnel"; Level means level of care, Volume means the number of VLBWIs per year, and Personnel means the number of neonatologists.
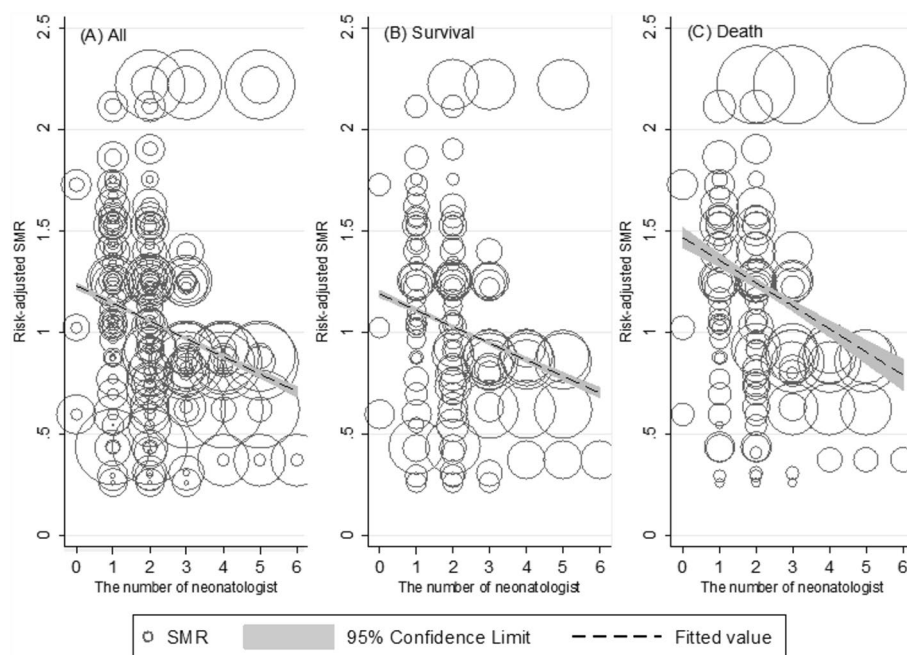

Figure 3. Patient risk adjusted SMR according to the number of neonatologists between survivors and non-survivors. These plots help us to visualize the association of the SMR on the number of neonatologists. The size of circles represents the number of VLBWIs as weight information. The result of testing the difference in regression slop across death and survival group identify as being statistically significant (*P*-value < .001).

The level of care is known to influence inter-hospital outcome variability[25,26,28,35,60], except in one large-scale study conducted in the United Kingdom[30].

On the contrary, several studies have highlighted the volume effect of admitted patients, which is considered one of the strongest hospital factors, with a higher volume associated with better outcomes owing to increased patient care experience and concentration[24–26,28]. One study demonstrated that the explanatory power of these factors in hospital-level variation reached up to 26%[26]. However, patient volume only partially influences inter-center variation[61], and conflicting reports exist, suggesting poorer outcomes in large units[62] or no significant impact[8,30]. This variability is likely attributable to differences in workload among healthcare professionals or uncontrolled factors in NICUs.

In contrast, studies on staffing are limited. The nurse-to-patient ratio[61] and nurse workload[63] in NICUs have been reported to affect outcomes. However, studies on the direct impact of the number of neonatologists are

scarce. In 1985, the American Academy of Pediatrics proposed an average of six VLBWIs per neonatologist, considering the severity, length of stay, and workload of NICU admissions[64]. Furthermore, based on population data and neonatal medical workforce status, in 2002 Thompson et al.[36] calculated the number of NICU beds (intensive and intermediate) per neonatologist as 8.4 in the United States, 6.9 in Australia, and 8.1 in Canada. In comparison, recent surveys in Korea showed an average of 13.4 NICU beds per neonatologist, which is almost twice as high as in Japan (7.0) or Taiwan (6.3)[38]. While this is a significant improvement compared with the 22.1 reported in a nationwide survey in 2006, few improvements have been made since surveys conducted in the 2010s. In the present study, the average number of NICU beds per neonatologist was 15.6 ± 7.3, and an average of 2.3 ± 1.1 neonatologists worked in Korean NICUs. Among the 15,766 patients, 66.3% were cared for in units with ≤ 2 neonatologists. However, evidence demonstrating that a lower patient load leads to better treatment outcomes is lacking. Although a large-scale study showed that regions with too few neonatologists per birth had an increase in neonatal mortality[32], no significant association between the number of neonatologists and neonatal survival rates has also been shown[30]. Nonetheless, the supply of neonatologists in Korea falls short of international standards, potentially leading to different analysis results compared to studies conducted in countries or institutions with adequate neonatological staffing. Additionally, NICU occupancy has been shown to affect mortality variation, highlighting the importance of neonatologist workloads[30]. In the present study, an increase in the number of neonatologists showed a clearer correlation with a decreased risk-adjusted SMR in the deceased group than in the surviving group. Furthermore, adjusting for hospital factors reduced the number of centers that exceeded the + 95% CI on the funnel plot, further supporting the finding that the number of working neonatologists influences inter-hospital mortality variation in Korea.

In this study, the absolute number of neonatologists working in a unit was selected as the more important variable than the number of beds per neonatologist. This may be attributed to the fact that only 13 of the included units had ≥ 3 neonatologists. Likewise, a previous analysis of NICUs in Korea[38] found no association between the number of beds per neonatologist and NICU level of care.

This study had some limitations. First, it focused on only three hospital factors as explanatory variables. This approach did not consider more detailed information, such as the composition of nursing and other physician workforces, care quality, and referral systems. In particular, the effect of the number of neonatologists on reducing inter-hospital SMR variation accounted for only 11.36% of the overall variation, whereas the level of care and patient volume showed no explanatory power. Therefore, further analysis is needed to explore other factors that might explain inter-hospital SMR variations in Korea.

Second, the number of neonatologists may not accurately represent the workload because factors such as the number of NICU beds per neonatologist are unaccounted. No official nurse practitioner system existed in Korea before 2020. Moreover, there was a significant shortage of fellows and pediatricians working in NICUs during the study period, with most of the workforce consisting of pediatric residents. Therefore, these results may reflect not only workload issues specific to neonatologists, but also the importance of neonatologist coverage within the NICU.

Despite these limitations, this nationwide study employed appropriate statistical modeling to accurately assess inter-hospital case-mix SMR variations and identified the number of working neonatologists as a significant hospital factor.

Meanwhile, it is important to note that there are concerns that comparative measures of mortality and morbidity are often overinterpreted, particularly in acute medical settings. This leads to unfair judgments about care quality and the potential stigmatization of institutions. Therefore, outcome data should be used for research and trend monitoring, while patient care improvements should focus on adherence to clinical and managerial standards[11].

In conclusion, the number of neonatologists was independently associated with center-to center-differences in VLBWI mortality in Korea after adjustment for patient risks and hospital factors. These findings provide important implications for conducting quality improvement initiatives in societies or countries with similar NICU systems, whereby expanding the number of neonatologists may be one of the key strategies for reducing mortality variation and improving outcomes for VLBWIs.

Data availability

References

1. Hack, M. *et al.* Very-low-birth-weight outcomes of the national institute of child health and human development neonatal network, november 1989 to october 1990. *Am. J. Obstet. Gynecol.* **172**, 457–464 (1995).
2. Stoll, B. J. *et al.* Trends in care practices, morbidity, and mortality of extremely preterm neonates, 1993–2012. *JAMA.* **314**, 1039–1051 (2015).
3. Vohr, B. R. *et al.* Center differences and outcomes of extremely low birth weight infants. *Pediatrics* **113**, 781–789 (2014).
4. Lee, S. K. *et al.* Variations in practice and outcomes in the canadian nicu network: 1996–1997. *Pediatrics* **106**, 1070–1079 (2000).

5. Tommiska, V. *et al.* A national short-term follow-up study of extremely low birth weight infants born in finland in 1996–1997. *Pediatrics* **107**, E2 (2001).
6. Berger, T. M. *et al.* Trends and centre-to-centre variability in survival rates of very preterm infants (<32 weeks) over a 10-year-period in switzerland. *Arch. Dis. Child. Fetal Neonatal Ed.* **97**, F323-328 (2012).
7. Kusuda, S. *et al.* Morbidity and mortality of infants with very low birth weight in japan: Center variation. *Pediatrics* **118**, e1130-1138 (2016).
8. Rogowski, J. A. *et al.* Indirect vs direct hospital quality indicators for very-low-birth-weight infants. *JAMA* **291**, 202–209 (2004).
9. Horbar, J. D., Badger, G. J., Lewit, E. M., Rogowski, J & Shiono, P. H. Hospital and patient characteristics associated with variation in 28-day mortality rates for very low birth weight infants Vermont oxford network. *Pediatrics* **99**, 149–156 (1997).
10. Simpson, J. M. *et al.* Analysing differences in clinical outcomes between hospitals. *Qual. Saf. Health Care.* **12**, 257–262 (2003).
11. Lilford, R., Mohammed, M. A., Spiegelhalter, D. & Thomson, R. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *Lancet.* **363**, 1147–1154 (2004).
12. Kleinman, J. C. Indirect standardization of neonatal mortality for birth weight. *Int. J. Epidemiol.* **11**, 146–154 (1982).
13. Organization WH. Neonatal and perinatal mortality : Country, regional and global estimates. 2006
14. Castro, E. C., Leite, A. J., Almeida, M. F. & Guinsburg, R. Perinatal factors associated with early neonatal deaths in very low birth weight preterm infants in northeast brazil. *BMC Pediatr.* **14**, 312 (2014).
15. Johansson, A. L., Dickman, P. W., Kramer, M. S. & Cnattingius, S. Maternal smoking and infant mortality: Does quitting smoking reduce the risk of infant death?. *Epidemiology.* **20**, 590–597 (2009).
16. Park, J. H., Chang, Y. S., Ahn, S. Y., Sung, S. I. & Park, W. S. Predicting mortality in extremely low birth weight infants: comparison between gestational age, birth weight, apgar score, crib ii score, initial and lowest serum albumin levels. *PLoS One.* **13**, e0192232 (2018).
17. Ghobad, M. *et al.* The relationship between maternal diseases during pregnancy and low birth weight: A nested case-control study in rural areas of kurdistan province (west of iran). *Int. J. Pediatr.* **5**, 5501–5514 (2017).
18. Sankaran, K. *et al.* Variations in mortality rates among canadian neonatal intensive care units. *Can. Med. Assoc. J.* **166**, 173–178 (2002).
19. Miyata, H. *et al.* Performance of in-hospital mortality prediction models for acute hospitalization: Hospital standardized mortality ratio in japan. *BMC Health Serv. Res.* **8**, 229 (2008).
20. Abdel-Latif, M. E., Nowak, G., Bajuk, B., Glass, K. & Harley, D. Variation in hospital mortality in an australian neonatal intensive care unit network. *Arch. Dis. Child. Fetal Neonatal Ed.* **103**, F331–F336 (2018).
21. Solomon, P. J., Kasza, J. & Moran, J. L. Australian and New Zealand intensive care society (ANZICS) centre for outcome and resource evaluation (CORE). Identifying unusual performance in Australian and New Zealand intensive care units from 2000 to 2010. *BMC Medical Res. Methodol.* **14**, 1–4 (2014).
22. Rosenblatt, R. A., Mayfield, J. A., Hart, L. G. & Baldwin, L. M. Outcomes of regionalized perinatal care in Washington state. *West. J. Med.* **149**, 98–102 (1988).
23. Mohammed, M. A., Manktelow, B. N. & Hofer, T. P. Comparison of four methods for deriving hospital standardised mortality ratios from a single hierarchical logistic regression model. *Stat. Methods Med. Res.* **25**, 706–715 (2016).
24. Bartels, D. B., Wypij, D., Wenzlaff, P., Dammann, O. & Poets, C. F. Hospital volume and neonatal mortality among very low birth weight infants. *Pediatrics* **117**, 2206–2214 (2006).
25. Phibbs, C. S. *et al.* Level and volume of neonatal intensive care and mortality in very-low-birth-weight infants. *N. Engl. J. Med.* **356**, 2165–2175 (2007).
26. Chung, J. H. *et al.* The effect of neonatal intensive care level and hospital volume on mortality of very low birth weight infants. *Med. Care.* **48**, 635–644 (2010).
27. Esser, M., Lack, N., Riedel, C., Mansmann, U. & von Kries, R. Relevance of hospital characteristics as performance indicators for treatment of very-low-birth-weight neonates. *Eur. J. Public Health.* **24**, 739–744 (2014).
28. Jensen, E. A. & Lorch, S. A. Effects of a birth hospital's neonatal intensive care unit level and annual volume of very low-birth-weight infant deliveries on morbidity and mortality. *JAMA Pediatr.* **169**, e151906 (2015).
29. Lee, H. C. *et al.* The impact of statistical choices on neonatal intensive care unit quality ratings based on nosocomial infection rates. *Arch. Pediatr. Adolesc. Med.* **165**, 429–434 (2011).
30. Tucker J. & Group UKNSS. Patient volume, staffing, and workload in relation to risk-adjusted outcomes in a random stratified sample of uk neonatal intensive care units: A prospective evaluation. *Lancet* **359**, 99–107 (2002).
31. Profit, J. *et al.* The association of level of care with nicu quality. *Pediatrics* **137**, e20144210 (2016).
32. Goodman, D. C. *et al.* The relation between the availability of neonatal intensive care and neonatal mortality. *N Engl J Med.* **346**, 1538–1544 (2002).
33. Yun, J. *et al.* Birth rate transition in the republic of Korea: Trends and prospects. *J. Korean Med. Sci.* **37**, e304 (2022).
34. Yum, H. K., Lim, C. H. & Park, J. Y. Medicosocial conflict and crisis due to illegal physician assistant system in korea. *J. Korean Med. Sci.* **36**, e199 (2021).
35. Shim, J. W. *et al.* The impact of neonatal care resources on regional variation in neonatal mortality among very low birthweight infants in korea. *Paediatr. Perinat. Epidemiol.* **27**, 216–225 (2013).
36. Thompson, L. A. & Goodman, D. C. Is more neonatal intensive care always better? Insights from a cross-national comparison of reproductive care. *Pediatrics* **109**, 1036–1043 (2002).
37. Chang, Y. S. Moving forward to improve safety and quality of neonatal intensive care in korea. *J Korean Med Sci* **33**, e89 (2018).
38. Lee, B. S., Lim, J. W., Choi, Y. S. & Kim, K. S. Current status of neonatologist staffing and workload in korean neonatal intensive care units. *Neonatal Med.* **27**, 65–72 (2020).
39. Chang, Y. S., Park, H. Y. & Park, W. S. The korean neonatal network: An overview. *J Korean Med Sci.* **30**(Suppl 1), S3–S11 (2015).
40. Gini R. & Forni S. Funnel plots for institutional comparisons. United kingdom stata users' group meetings 2009. (2009).
41. Horbar, J. D. The vermont oxford network: evidence-based quality improvement for neonatology. *Pediatrics* **103**, 350–359 (1999).
42. Seaton, S. E., Barker, L., Lingsma, H. F., Steyerberg, E. W. & Manktelow, B. N. What is the probability of detecting poorly performing hospitals using funnel plots?. *BMJ Qual Saf.* **22**, 870–876 (2013).
43. Stark A. R. & American Academy of Pediatrics Committee on F, Newborn. Levels of neonatal care. *Pediatrics.* 114, 1341–1347 (2004).
44. Zhang, Z. Variable selection with stepwise and best subset approaches. *Ann. Transl. Med.* **4**, 136 (2016).
45. Schwarz, G. Estimating the dimenstion of a model. *Ann Statist.* **6**, 461–464 (1978).
46. Akaike H. On entropy maximization principle. Aplication of statistics. (1977)
47. F Harrell (2001) Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis. Springer
48. Lemeshow, S. & Hosmer, D. W. Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am. J. Epidemiol.* **115**(1), 92–106 (1982).
49. EW Steyerberg et al. 2001. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. J. Clin. Epidemiol. 54. 774–781
50. HFE Jr KL. Lee DB Mark. 1996. Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat. Med. 15. 361–387

51. Spiegelhalter, D. J. Funnel plots for comparing institutional performance. *Stat. Med.* **24**, 185–1202 (2005).
52. Agresti A. *Categorical data analysis (3rd ed)*. Hoboken, New Jersey: John Wiley & Sons; 2013.
53. Leyland A.H. & Groenewegen, P. P. *Multilevel modelling for public health and health services research: Health in context*. [S. I.]: Springer International Publishing; 2020.
54. Snijders, T. B. *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (Sage Publications, 1999).
55. Merlo, J. *et al.* A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *J. Epidemiol. Community Health.* **60**, 290–297 (2006).
56. Gupta, R. C. Estimation of standardized mortality ratio with missing death certificates. *Math Comput Model.* **40**, 491–498 (2004).
57. Moran J.L. & Solomon P. J. (2014). Outcome ACf, Resource Evaluation of A, New Zealand Intensive Care S. Fixed effects modelling for provider mortality outcomes: Analysis of the australia and new zealand intensive care society (anzics) adult patient data-base. *PLoS One.* 9, e102297
58. Matsui, H., Fushimi, K. & Yasunaga, H. Variation in risk-standardized mortality of stroke among hospitals in japan. *PLoS ONE* **10**, e0139216 (2015).
59. MacKenzie, T. A. *et al.* A primer on using shrinkage to compare in-hospital mortality between centers. *Ann Thorac Surg.* **99**, 757–761 (2015).
60. Lapcharoensap, W. *et al.* Hospital variation and risk factors for bronchopulmonary dysplasia in a population-based cohort. *JAMA Pediatr.* **169**, e143676 (2015).
61. Sherenian, M. *et al.* Nurse-to-patient ratios and neonatal outcomes: A brief systematic review. *Neonatology* **104**, 179–183 (2013).
62. Shah, P. S., Mirea, L., Ng, E., Solimano, A. & Lee, S. K. Canadian Neonatal N. Association of unit size, resource utilization and occupancy with outcomes of preterm infants. *J Perinatol.* **35**, 522–529 (2015).
63. Tubbs-Cooley, H. L., Mara, C. A., Carle, A. C., Mark, B. A. & Pickler, R. H. Association of nurse workload with missed nursing care in the neonatal intensive care unit. *JAMA Pediatr.* **173**, 44–51 (2019).
64. American academy of pediatrics. Committee on fetus and newborn. *Manpower needs in neonatal pediatrics. Pediatrics.* **76**, 132–135 (1985).

Acknowledgements

Author contributions

Myung Hee Lee: conceptualization, data curation, formal analysis, methodology, writing Jang Hoon Lee: conceptualization, formal analysis, methodology, writing, editing Yun Sil Chang: conceptualization, formal analysis, funding acquisition, methodology, project administration, supervision, writing. # Myung Hee Lee and Jang Hoon Lee contributed equally as co-first authors.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-69680-1.

Correspondence and requests for materials should be addressed to Y.S.C.

Reprints and permissions information is available at www.nature.com/reprints.