

Developing a streamlined risk-adjusted cesarean section rate model for evaluation of obstetrical quality across hospitals by using EHRs: A provincial-scale multicenter retrospective study

DIGITAL HEALTH
Volume 10: 1–17
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076241284726
journals.sagepub.com/home/dhj



Da Zhou^{1,*}, Wanting Zhong^{3,*}, Qiu Sun^{4,*}, Qiang Fu⁵, Pei Liu², Shilin Zhong², Guoqing Li⁶, Bin Luo⁷, Xiao Chen⁸, Jian Wang⁹ and Chang Xu² 

Abstract

Objective: This study aims to explore a streamlined risk-adjusted cesarean section rate (RCSR) model and to compare its practical application effects with the traditional RCSR models.

Methods: Utilizing obstetric electronic health record (EHR) data from provincial multicenter hospitals, this study establishes a streamlined RCSR model alongside the traditional RCSR model and evaluates the efficacy of both models. Subsequently, the RCSRs of 56 hospitals within the province are calculated and ranked using both models. The consistency of these rankings is then quantified using Kendall's tau coefficient of concordance.

Result: Comparison of model effectiveness evaluation of the traditional RCSR model versus the streamlined RCSR model is as follows: AUC (0.840 vs 0.839), accuracy (0.875 vs 0.872), sensitivity (0.690 vs 0.685), specificity (0.898 vs 0.892), positive predictive value (0.908 vs 0.903), negative predictive value (0.664 vs 0.660), and Brier score (0.069 vs 0.067). In the test of the consistency of hospital rankings based on two models, Kendall's tau coefficients were observed to be 0.979 (year 2017), 0.978 (year 2018), and 0.978 (year 2019) over a span of 3 years, with an aggregate coefficient of 0.974.

Conclusion: In the realm of model performance evaluation as well as the pragmatic application within hospital settings, the streamlined model exhibits a substantial congruence with the traditional model. Therefore, the streamlined model can effectively serve as a viable surrogate for the traditional model, potentially establishing itself as a refined paradigm for the appraisal of quality in obstetric healthcare services.

Keywords

Risk adjustment, cesarean section rate, obstetrical quality evaluation, streamlined modeling, EHRs

Submission date: 5 April 2024; Acceptance date: 16 August 2024

¹Huazhong University of Science and Technology, School of Public Health, Tongji Medical College, Wuhan, China

²Peking University Shenzhen Hospital, Shenzhen, China

³Zhuhai People's Hospital (Zhuhai Hospital Affiliated with Jinan University), Zhuhai, China

⁴Wuhan No.1 Hospital, Wuhan, China

⁵Tufts University, Medford, MA, USA

⁶Beijing Jishuitan Hospital Affiliated to Capital Medical University, Beijing, China

⁷Wuhan Diaolong Data Technology Co., Ltd., Wuhan, China

⁸The University of British Columbia–Vancouver Campus, Vancouver, Canada

⁹Wuhan Children's Hospital, Wuhan, China

*DZ, WZ, and QS contributed equally to this paper as the co-first authors.

Corresponding authors:

Chang Xu, Peking University Shenzhen Hospital, No. 1120, Lianhua Road, Futian District, Shenzhen, Guangdong Province 518036, China.

Email: wanchang0421@sina.com

Jian Wang, Wuhan Children's Hospital, Wuhan, China.

Email: baal76@sina.com



Introduction

Compared with other clinical departments in hospitals, obstetrics has the highest incidence of medical errors and medical malpractice claims. A recent study has reported that obstetrics contributes to 10% of all medical malpractice claims in hospitals, 76% of which involve medical errors.¹ The operation of obstetrics affects the health and life safety of parturients and newborns, who are often accompanied by high medical risks due to complex individual differences in parturients, intrapartum complications, and a variety of unpredictable outcomes.^{2–6} Consequently, pursuing research focusing on the quality and safety of obstetrical care presents both a critical need and a formidable challenge.

As the front-end link of obstetrical quality improvement, obstetrical quality assessment is getting more research attention than ever before. Many obstetric experts have been trying to detect the best obstetrical medical quality indicators and evaluation methods^{4,5,7–16}; however, a universal consensus remains elusive, until Bailit, through a series of multicenter studies, proposed that an exemplary obstetrical quality metric for inpatient settings should embody five principal attributes^{7–12}: (1) association with meaningful maternal and neonatal outcomes, (2) relation to outcomes that are influenced by physician/health system behaviors, (3) affordability for application on a large-scale basis, (4) acceptability to practicing obstetricians as a meaningful marker of quality, and (5) reliability/reproducibility. Meanwhile, Bailit advocated for the risk-adjusted cesarean section rate (RCSR) as the quintessential index, attributing its recommendation to simultaneously fulfilling the aforementioned five characteristics concerning obstetrical quality.¹⁷

With the advancement of hospital informatization, it has been observed that the individual variances among parturients can be comprehensively captured within their electronic health records (EHRs).^{2,3} This phenomenon provides a foundation for the refinement of the RCSR model and facilitates a more nuanced re-ranking of hospitals after adjusting the risk factors at the parturient level. Nevertheless, the execution of research related to risk adjustment mandates the procurement of extensive multicenter medical datasets and the application of relatively sophisticated statistical analysis techniques.¹² These requirements (multicenter, EHRs big data, sophisticated statistical analysis) have led to a scarcity of high-caliber studies and reports, particularly emanating from China, in this domain over the past two decades. In addition, even the limited studies focused on cesarean section risk adjustment models have predominantly utilized almost total indications for cesarean sections as independent variables in model formulation.^{18,19} Such an approach frequently results in an excessive number of variables, which, in turn, can lead to increased model complexity and issues such as overfitting.²⁰ Therefore, developing a streamlined RCSR model that maintains the effectiveness of traditional

RCSR models holds significant importance for the global assessment of obstetric medical quality.

This research endeavors to undertake a rigorous empirical analysis of the quality of obstetrical care by assessing the RCSR through data derived from the electronic medical records (EMRs) of 56 tertiary-level public hospitals in Hubei Province. This study aspires to enhance and stratify parturients' individual characteristics based on the extensive review and synthesis of extant scholarship. The resultant streamlined risk-adjusted cesarean section model will be methodically compared with established traditional models. Such a comparison is poised to constitute the crux of this study, promising to contribute to refining the quality assessment model within obstetrical practice.

Methods

Research data sources

This study was supported by the Health Commission of Hubei Province at the beginning of the design. The study used the de-identified maternal EMRs from the obstetrical discharge database. The database included the following information about each parturient: demographic characteristics (maternal age, gestational age, hospitalization, pregnancy and birth history, newborn weight, and obstetrical complications), all diagnostic descriptions, and the International Classification of Diseases (ICD-10) codes. Each record included up to 10 secondary diagnoses and their ICD-10 codes, up to seven medical procedures and their ICD-9-CM operation codes, and maternal discharge outcomes.

The data collection process of the database was described as follows: Firstly, a physician reviewed every case in the hospital. Secondly, trained coders certified by the Medical Record Management Association (MRMA) entered the patients' data into the administrative system. Lastly, a hospital case manager cross-checked the data entry before uploading it to the EMR database to ensure accuracy and quality. This quality assurance program reduced human errors in data entry.

Main study design

We mainly designed an empirical comparative study: The traditional cesarean section risk adjustment model (full factor) and the streamlined cesarean section model (optimized classification of obstetrical complications) were used to analyze and evaluate the data of obstetrical EMRs from 2017 to 2019 in 56 tertiary public hospitals in Hubei Province which had obtained Chinese grading and evaluation of EMRs of more than level 3 through the consistency test to compare the similarity of the adjustment results. This study aimed to measure whether the practicability and stability of the streamlined model could reach the effectiveness of the traditional model.

Patient and public involvement

This study is a medical and health academic research carried out by deleting patients' sensitive information (ID code, name, phone number, address, etc.) with the permission of the Information Center of the Hubei Provincial Health Commission. We informed parturients through electronic messages that their production information would be used for quality evaluation among medical institutions after deleting sensitive information. There is no information leakage in the process, which strictly ensures the security of patient information.

Samples and inclusion/exclusion criteria

We selected 56 large tertiary public hospitals in various local regions and municipalities of Hubei Province for this study. A total of 467,521 electronic health records of parturients were obtained from 1 January 2017 to 31 December 2019. The inclusion and exclusion criteria were adopted from the American College of Obstetricians and Gynecologists (ACOG),^{12,21} and we made appropriate adjustments according to the actual situation in China. The inclusion criteria were as follows: (1) maternal age between 16 years and 49 years, inclusive; (2) gestational age greater than 23 weeks; and (3) delivery of at least one fetus that survived before discharge. The exclusion criteria were as follows: (1) delivery for non-reproductive purposes (such as abortion and induced labor. These situations would be much different from reproductive purpose labor in fetal progress); (2) loss of core maternal variables (such as obstetrical complications or delivery mode information); and (3) illogical data error or rare cases (such as birth weight of less than 500 g). The final number of discharge records of

eligible parturients who met all the criteria was 391,661. The specific inclusion and exclusion criteria and the screening process are shown in Figure 1.

It should be noted that in ACOG obstetric medical quality evaluation standard, puerpera with "normal fertility willingness" and in-hospital delivery are the qualified records by default. (There is a significant difference in the standardized medical service content between pregnant women with and without fertility intentions, so the two do not have homogeneity in the evaluation of obstetric medical quality). Therefore, women who are too young (<16 years old), and those who intended to have abortions or induced labor were excluded from the sample of this study.

Main outcome

Whether the parturient underwent CS during hospitalization or not is the outcome variable. According to her electronic record, we used the ICD procedures and operation codes (ICD-9-CM3) to identify whether the parturient had a CS during hospital delivery. The ICD-9-CM Volume 3 procedures and operation codes corresponding to CS included the following: 74.000, 74.100, 74.101, 74.102, 74.200, 74.400, 74.401, and 74.9900.

The variation composition of two RCSR models

The variation of the traditional RCSR model and the streamlined RCSR model is mainly composed of three parts: (1) maternal basic characteristics (maternal age, maternal registration status upon admission, previous cesarean, gestational age, neonatal birth weight), (2) pregnancy complications (fetal position, cephalopelvic disproportion, umbilical cord

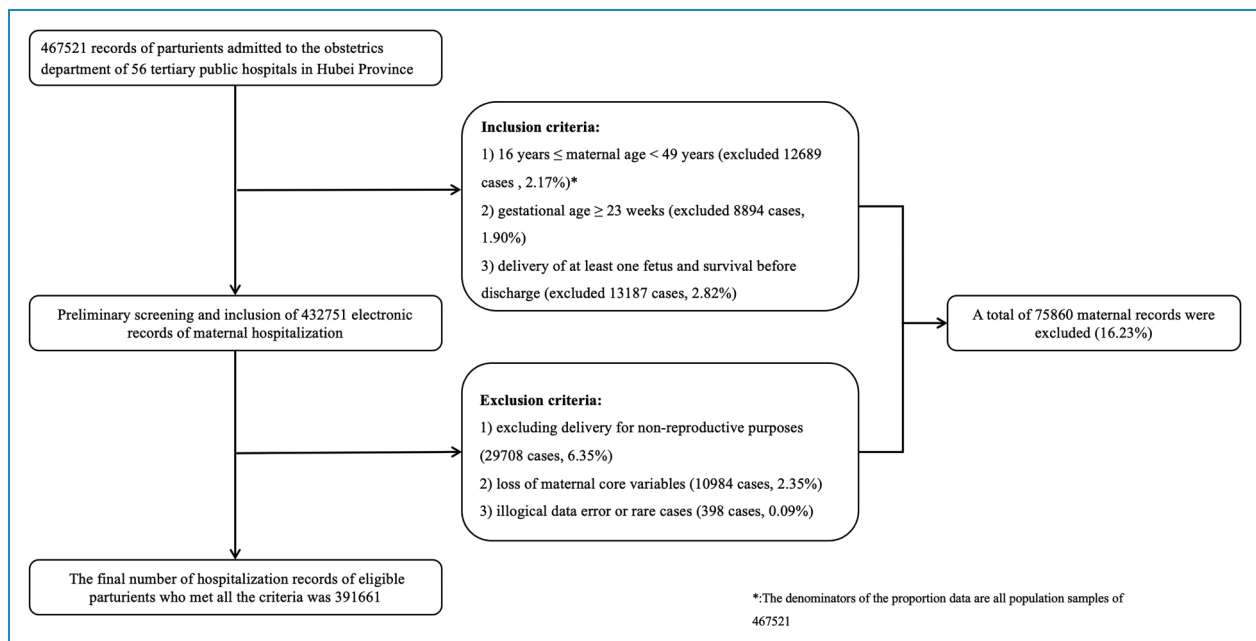


Figure 1. Flowchart of the sampling process.

entanglement, placenta previa, placental implantation and adhesion, placental abruption, premature rupture of membranes, situation of labor process), and (3) medical conditions of pregnancy (incompetent cervix, preeclampsia, any hypertension, diabetes mellitus, cardiac diseases, acute or chronic lung disease, hepatitis B, intrahepatic cholestasis, thyroid dysfunction, syphilis, Rh sensitization). Related diagnoses were also selected according to the ICD-10 codes of the parturients' medical records. (The specific diagnostic codes are shown in Supplemental Table 1).

The difference between the two models mainly lies in the third part: We screened the medical conditions of pregnancy related to the cesarean section and have integrated the opinions of classification of pregnancy complications by combining the three parties—ACOG standards,²¹ Bailit series of studies,^{10–12} and specification for risk assessment and management of parturient from the National Health Commission of China.²² Then as the core unit of the streamlined RCSR model, the variables in the third part were divided into four levels, mild, moderate, severe, and none, as shown in Figure 2. In addition, for parturients with multiple different levels of medical conditions, the diagnosis at the highest level would be taken as the basis for classification.

Statistical analysis

Primary model development. In order to make the model simple and feasible, we classified the continuous variables in both two RCSR models. The Chi-square test and single-factor logistic regression were used to examine the association between CS and each independent variable. Whatever model, those significant variables ($P < 0.05$)

were included in the subsequent risk adjustment model unless supported by obstetric clinical experts. The variance inflation factor (VIF) was used to test the collinearity.²³ The variables that passed the above two steps were included in the subsequent cross-validation analysis.

Main model design: hierarchical-multivariate logistic regression model based on cross-validation. We made a deep combination of hierarchical-multivariate model and cross-validation to solve the problem of parturient aggregation which happens in multi-medical center research and model training at the same time. First, we considered the characteristic differences of maternal individuals as fixed effects and the events of maternal visits to any hospital as random effects, to establish a hierarchical multiple logistic models of these two effects. This type of model would be applied to the actual risk adjustment process in both traditional and streamlined model operations. Second, we used the k -fold cross-validation method to create a derivation and a verification data set.¹² In order to ensure the fairness and effectiveness of cross-validation for every hospital, all samples were randomly divided into 10 approximately equal size groups by the unit of each hospital as 10 sub-samples (e.g. in a hospital with an annual hospitalization capacity of 1000, we require that 100 parturients should be randomly included in the cross-validation sub-grouping process). Hierarchical-multivariate logistic regression was applied using any nine possible combinations of 10 sub-samples. The stepwise method was used to retain the variables with $P < 0.05$. The remaining one sub-sample of the 10 validation datasets was used to validate. The results of the validation were evaluated using the C-statistic. The C-statistic was calculated by the area value under the

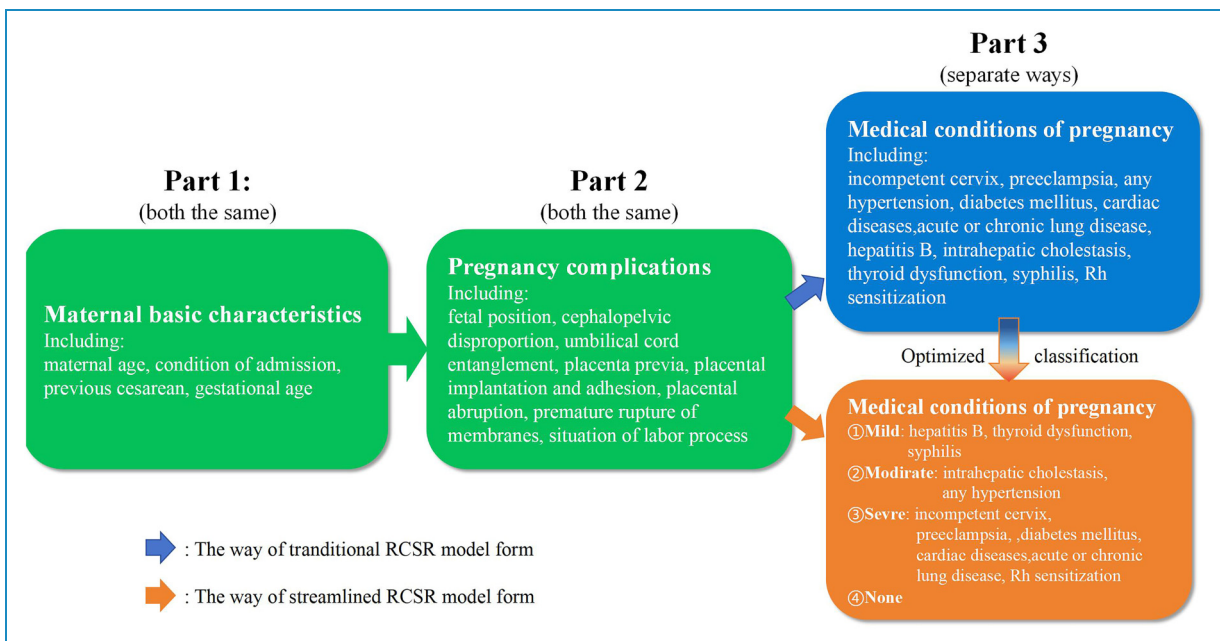


Figure 2. The variation composition of two RCSR models.

curve (AUC) of the receiver operating characteristic (ROC). The verification results were obtained 10 times. Then the model constructed by the corresponding nine sub-samples under the verification of the highest C-statistics (while each variable in the top C-statistics model needs to appear at least eight times in any other sub-sample model¹²) is taken as a training set and as the final prediction model of risk adjustment, which is used to calculate the expected probability of all samples. The odds ratios (OR) and 95% confidence intervals (CI) were reported.

Evaluating model performance. The performance of the traditional and streamlined models was mainly assessed by its differentiation (AUC of ROC).²⁴ Other model evaluation indicators are observed as additional references, such as accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and Brier score.²⁵ Usually, the calibration of modeling was used to reflect the goodness-of-fit of the model, and the Hosmer-Lemeshow method was often used for the testing. Unfortunately, some scholars have clearly pointed out that a sample size as large as this study is unsuitable for using this method for evaluation.²⁶

RCSR calculating and obstetric ranking. The hierarchical multiple logistic regression model is written as follows:

$$\text{Logit}(P) = (y_{ij} = 1|X_{ij}) = \alpha + \beta X_{ij} \quad (1)$$

The parameter y_{ij} is the outcome variable of whether the parturient used CS for delivery during hospitalization, which is equivalent to the occurrence of the index outcome of the j parturient in the obstetrics department of the i hospital. It is usually represented by 0 (marked as not happening) and 1 (marked as occurring), respectively. X_{ij} contains the characteristic variables of the parturient (including maternal age, condition of admission, gestational age, birth weight, and obstetrical complications). According to the formula, the parameters α and β are considered as fixed effects, and the function connection uses the transformation method of the logit as the link function.

In addition, the parameter α also needs to be satisfied by the following formula:

$$\alpha = \mu + \omega_i, \quad \omega_i \sim N(0, \tau^2) \quad (2)$$

ω_i is the random effect, and the larger τ^2 is, the stronger the aggregation of data is. When τ^2 is 0, the model evolves into a general logistic regression model.

After estimating the actual parameters $\hat{\alpha}$ and $\hat{\beta}$, the probability of model-based CSR of the j parturient who was actually hospitalized in the i obstetric could be obtained as follows:

$$\hat{p}_{ij} = \frac{\exp(\hat{\alpha} + \hat{\beta}X_{ij})}{1 + \exp(\hat{\alpha} + \hat{\beta}X_{ij})} = L^{-1}(\hat{\alpha} + \hat{\beta}X_{ij}) \quad (3)$$

The observed and expected CS of all parturients in the

obstetrics department were compared, and the standard rate (SR) of CS was calculated. The ratio of the total number of parturients actually observed in the obstetrics department at the total number of expected CS (observed/expected, O/E) was calculated. Then the formula for the incidence of the standardized obstetrical outcome in the obstetrics department was described as follows:

$$SR_i = \left(\sum_{j=1}^{n_i} y_{ij} / \sum_{j=1}^{n_i} \hat{p}_{ij} \right) \times p_{\mu} \quad (4)$$

Obstetrics in all hospitals were ranked according to the traditional RCSR evaluation outcomes and the streamlined RCSR outcomes. The agreement of obstetrics rankings between the two assessment methods was tested using Kendall's coefficient of concordance test. A significant difference level was set at $P < 0.05$ (two_tailed). All of the analyses were conducted using STATA 14.0.²⁷

Results

Descriptive result

Hospital/obstetric basic situation. The descriptive statistics of CS in 56 obstetrics of tertiary public hospitals from 2017 to 2019 are shown in Supplemental Table 2. The overall CSR was 59.34%, and the annual average CSR was 58.38% (2017), 59.89% (2018), and 59.78% (2019). It is reiterated here that the statistical analysis of the cesarean section rate (CSR) in this study was conducted based on the inclusion and exclusion criteria specified in this study (excluding EMRs for non-reproductive purposes and those with missing core data). This might result in a higher CSR compared to the conventional statistical methods.

There are 18 provincial capital hospitals (32.14%) and 38 non-provincial capital hospitals (67.86%). There are 47 teaching hospitals (83.93%) and 9 non-teaching hospitals (16.07%). There are 52 comprehensive hospitals (92.86%), 3 maternal and child healthcare hospitals (5.36%), and 1 traditional Chinese medicine hospital (1.79%).

Parturient basic situation. EMRs of 391,661 parturients were included in the final data analysis. The average age and standard deviation of parturients were 28.99 (y)±5.42 (y). The average gestational age was 38 (w)±2.72 (d), while the standard deviation was 21.40 (d). The average birth weight and standard deviation of newborn babies were 3232.98 (g)±504.41 (g). The distributions of maternal characteristics are shown in Table 1.

Obstetrical epidemiology content comparison for two models

Chi-square test. Table 1 exhibits the Chi-square test results for variables. We noticed that no significant differences were found for Rh sensitivity. Moreover, as this variable did not pass cross-validation (less than eight times occurrences in sub-samples), it will also be removed from the final model.

Collinear analysis. According to Table 1, the VIF of most variables is below 2. This means the collinearity between these variables has almost no substantial effect on the model. However, we found the VIF of gestational age was > 10 (12.68). Its effect overlapped with the neonatal birth weight (8.86). Considering that neonatal birth weight is more likely to be an index of obstetrical outcome, not a predictor, and not entirely a maternal characteristic, the variable of neonatal birth weight was removed in the subsequent construction of the two models. After removal, the VIF of the gestational age variable was reduced to 1.17.

Logistic regression analysis. We conducted three logistic regressions for analysis: univariate, multivariate, and hierarchical-multivariate (as shown in Table 2). On this basis, we also added a comparison of OR values and CI between the traditional and streamlined models in two sets of multivariate regression analysis, in order to fully clarify the differences between different models in obstetric clinical practice. From the overall regression analysis results, there is a certain difference in the connotation between univariate and multivariate regression models. At the same time, there is a tiny difference between multivariate and hierarchical-multivariate regression models. In addition, it has also very little difference in the expression of the same part (maternal basic characteristics part and pregnancy complications part) between the traditional and streamlined models.

Cross-validation modeling exhibition

By comparing the graphs, we found that the two sets of models also have a high degree of similarity in cross-validation, even including the AUC area of the models (mean CV vs AUC: 0.835 vs 0.833) and the routing of the ROC curve (Figure 3).

Comparison of modeling effectiveness

The comparison of the performance evaluation between the traditional and streamlined models is shown in Table 3. It should be noted that both model performance evaluations were conducted using the samples corresponding to the highest AUC validation set during the cross-validation

process as the training set. Therefore, they were both constructed using the hierarchical-multivariate logistic regression model. The comparative data in Table 3 indicated that the two models are almost equivalent in terms of performance evaluation in various dimensions.

Consistency test results of obstetrical RCSR ranking

Figure 4 provides a detailed comparison of the changes in hospital rankings between two RCSR models over different years. In these graphs, the horizontal axis represents the RCSR ranking under the traditional RCSR model, while the vertical axis represents the streamlined version. Therefore, the closer the origin representing the hospital is to the diagonal in the graph, the smaller the change in ranking. It could be seen that the ranking of most hospitals has not undergone significant changes. Even in hospitals that have changed, the magnitude of their ranking changes is still relatively small. Meanwhile, Kendall's tau coefficients presented in Table 4 also confirmed this phenomenon.

Discussion

Cesarean section rates

CSR, as a pivotal metric for assessing the caliber of obstetric care, has been entrenched in global practice for numerous decades, as evidenced by a plethora of studies.^{7,17,28–30} Scholarly discourse in obstetrics has consistently indicated that an aberrant CSR, whether excessively high or low, precipitates a spectrum of negative obstetric sequelae.^{7,8,29} Thus, it is imperative to identify and uphold an optimal CSR range. The World Health Organization (WHO), in its 2015 discourse on CSRs, reaffirmed a longstanding international health community consensus dating back to 1985, which postulates that a CSR between 10% and 15%³¹ is most propitious, later suggesting a modest recalibration toward approximately 19%.³² Despite these guidelines, the optimal CSR warrants nuanced calibration in response to the heterogeneity of health statuses, disease profiles, cultural frameworks, economic conditions, and other regional variables, as advanced by scholars researching risk adjustment for public health.³³

In the course of the past three decades, China has witnessed a precipitous augmentation in the rate of cesarean deliveries.³⁴ An authoritative report shows that from 2008 to 2014, China's CSR increased from 28.8% to 34.9% and then to 36.7% in 2018, ranking first among Asian countries.³⁵ In 2010, the World Health Organization articulated critique concerning the disproportionately high CSRs within China, with the aspiration of catalyzing heightened attention and the implementation of efficacious measures by Chinese health regulatory authorities.^{36,37}

Table 1. The distributions of characteristics about 391661 maternal records.

Maternal characteristics	Classification or subgroup	Count (N)	Proportion (%)	Count of cesarean (N)	Cesarean rate (%)	χ^2	P-value	VIF
Maternal age	16–20 y	5578	1.42	2194	39.33	16124.57	< 0.001**	1.07
	21–25 y	69,696	17.80	32882	47.18			
	26–30 y	18,1887	46.44	100843	55.44			
	31–35 y	94,184	24.05	65308	69.34			
	> 35 y	40,316	10.29	31172	77.32			
Maternal registration status upon admission	Normal	10,9508	27.96	62781	57.33	1107.97	< 0.001**	1.21
	Emergency	35,592	9.09	18818	52.87			
	serious	24,6561	62.95	150800	61.16			
Previous cesarean	No	29,0184	74.09	133533	46.02	8174.60	< 0.001**	1.09
	Yes	10,1477	25.91	98866	97.43			
Number of fetuses	Singleton	38,2707	97.71	224164	58.57	4042.18	< 0.001**	1.05
	Twins or multiple births	8954	2.29	8235	91.97			
Gestational age	< 37 w	35,375	9.03	25499	72.08	3237.41	< 0.001**	12.68 Δ
	≥ 37 w while < 42 w	33,7573	86.19	192339	56.98			
	≥ 42 w	18,713	4.78	14561	77.81			
Neonatal birth weight	500–2499 g	25,617	6.54	18271	71.32	3078.78	< 0.001**	8.86
	2500–3999 g	34,5173	88.13	199174	57.70			
	≥ 4000 g	20,871	5.33	14954	71.65			
Fetal position	Normal	37,0172	94.51	212331	57.36	1091.00	< 0.001**	1.01
	Abnormal head position	6082	1.55	5500	90.43			
	Breech or transverse position	15,407	3.93	14,568	94.55			
Cephalopelvic disproportion	No	37,8288	96.59	219,129	57.93	9144.22	< 0.001**	1.02
	Yes	13,373	3.41	13,270	99.23			
	No	30,2201	77.16	176,872	58.53	358.74	< 0.001**	1.01

(continued)

Table 1. Continued.

Maternal characteristics	Classification or subgroup	Count (N)	Proportion (%)	Count of cesarean (N)	Cesarean rate (%)	χ^2	P-value	VIF
Umbilical cord entanglement	Yes	89,460	22.84	55,527	62.07			
Placenta previa	No	37,9127	96.80	220,380	58.13	7197.45	< 0.001**	1.06
	Yes	12,534	3.20	12,019	95.89			
Placental implantation and adhesion	No	36,4571	93.08	212,583	58.31	2314.56	< 0.001**	1.05
	Yes	27,090	6.92	19,816	73.15			
Placental abruption	No	38,8720	99.25	229,975	59.16	654.42	< 0.001**	1.01
	Yes	2941	0.75	2424	82.42			
Premature rupture of membranes	No	32,3308	82.55	198,661	61.45	3386.73	< 0.001**	1.02
	Yes	68,353	17.45	33,738	49.36			
Situation of labor process	Normal	36,1182	92.22	213,890	59.22	2832	< 0.001**	1.06
	Precipitate labor	1822	0.47	8	0.44			
	Prolonged labor process	1421	0.36	1102	77.55			
Medical conditions level of pregnancy	None	28,4701	72.69	159,548	56.04	5403	< 0.001**	1.28 [#]
	Mild	11,789	3.01	6954	58.99			
	Moderate	14,197	3.62	10,549	74.30			
	Severe	53,738	13.72	37,949	70.62			
Incompetent cervix	No	38,5762	98.49	227,816	59.06	836.22	< 0.001**	1.01
	Yes	5899	1.51	4583	77.69			
Preeclampsia	No	38,6223	98.61	228,324	59.12	556.12	< 0.001**	1.89
	Yes	5438	1.39	4075	74.94			
Any hypertension	No	37,4646	95.66	219,345	58.55	2239.00	< 0.001**	1.67
	Yes	17,015	4.34	13,054	76.72			
Diabetes mellitus	No	34,6449	88.46	200,835	57.97	2307.85	< 0.001**	1.71
	Yes	45,212	11.54	31,564	69.81			
Cardiac diseases	No	38,9764	99.52	230,846	59.23	401.00	< 0.001**	1.01

(continued)

Table 1. Continued.

Maternal characteristics	Classification or subgroup	Count (N)	Proportion (%)	Count of cesarean (N)	Cesarean rate (%)	χ^2	P-value	VIF
	Yes	1897	0.48	1553	81.87			
Acute or chronic lung disease	No	39,0493	99.70	231,656	59.32	8.88	0.003**	1.01
	Yes	1168	0.30	743	63.61			
Hepatitis B	No	38,9800	99.53	231,206	59.31	17.62	< 0.001**	1.01
	Yes	1861	0.48	1193	64.11			
Intrahepatic cholestasis	No	38,5320	98.38	227,375	59.01	1072.43	< 0.001**	1.01
	Yes	6341	1.62	5024	79.23			
Thyroid dysfunction	No	37,6068	96.02	222,617	59.20	77.65	< 0.001**	1.06
	Yes	15,593	3.98	9782	62.73			
Syphilis	No	39,0768	99.77	231,785	59.32	32.92	< 0.001**	1.01
	Yes	893	0.23	614	68.76			
Rh sensitization	No	39,1301	99.91	232,178	59.34	0.63	0.428	1.01
	Yes	360	0.09	221	61.39			
	Yes	650	0.17	354	54.46			

* $P < 0.05$; ** $P < 0.01$; Δ VIF > 10 .

#The VIF value of the variable “medical conditions level of pregnancy” was calculated based on the streamlined model.

Within the description analysis of our study, the collective CSR was registered at 59.34%, a statistic is directly linked to the stringent inclusion and exclusion criteria underpin a methodical evaluation of obstetric medical quality. Notwithstanding the non-adherence to this delineated treatment protocol, the CSR would remain at a disconcerting 49.71%, hence maintaining a considerable elevation above the idealized benchmark. Furthermore, the original CSR of each hospital encompassed within the sample exceeded 40%, substantially overstepping the thresholds advocated by the WHO. Consequently, in the realm of this investigation, CSR has been construed as an inferiority metric (lower values are indicative of better outcomes), necessitating a critical appraisal and strategic interventions to ameliorate with internationally endorsed standards.

Risk-adjusted CSR models

The meaning of RCSR in obstetrical management practice.

The import of the RCSR resides in its meticulous

contemplation of maternal determinants, alongside the abrogation of their influence upon cesarean section incidence through sophisticated risk-adjusted frameworks. Such models markedly enhance the capacity for an equitable comparison of the intrinsic quality of obstetric care proffered by disparate medical institutions.^{2,3,12} This aspect is quintessential for formulating and promulgating strategic initiatives designed to modulate the cesarean section frequency within a rational purview.²⁸ This is particularly germane to China, a nation that has grappled with inflated cesarean delivery rates in the preceding decennium, accompanying this trend is the shift from a “one-child policy” to a “two-child policy”^{33,34} (even the three-child policy right now).

An authoritative research has revealed a substantial variability in cesarean section frequencies across Chinese hospitals, a phenomenon more attributable to institutional factors (such as hospital management) than maternal characteristics.³⁸ This perspective aptly corroborates the significance of risk-adjusted obstetric care quality rankings. Although maternal factors may not be the primary

Table 2. The odds ratio (OR) and 95% confidence interval (CI) among logistic regressions.

Variables/Logistic regressions	Multivariate			Hierarchical-multivariate		
	Univariate	Traditional		Streamlined	Traditional	Streamlined
		Traditional	Streamlined			
Maternal age						
16-20 y	0.52 (0.49-0.55)**	0.82 (0.78-0.88)**	0.84 (0.79-0.89)**	0.84 (0.79-0.90)**	0.85 (0.80-0.91)**	
21-25 y	0.72 (0.70-0.73)**	0.95 (0.93-0.97)**	0.96 (0.94-0.98)**	0.95 (0.93-0.97)**	0.95 (0.93-0.97)**	
26-30 y	(Reference group)	(Reference group)	(Reference group)	(Reference group)	(Reference group)	
31-35 y	1.83 (1.80-1.86)**	1.04 (1.01-1.06)**	1.04 (1.02-1.06)**	1.04 (1.01-1.05)**	1.03 (1.01-1.06)**	
> 35 y	2.74 (2.67-2.82)**	1.40 (1.36-1.45)**	1.40 (1.36-1.45)**	1.38 (1.34-1.43)**	1.38 (1.34-1.42)**	
Maternal registration status upon admission						
Normal	(Reference group)	(Reference group)	(Reference group)	(Reference group)	(Reference group)	
Emergency	0.79 (0.77-0.81)**	0.81 (0.79-0.84)**	0.82 (0.79-0.84)**	0.72 (0.69-0.75)**	0.73 (0.70-0.76)**	
Relatively serious	1.17 (1.16-1.19)**	1.08 (1.06-1.11)**	1.09 (1.07-1.11)**	0.91 (0.88-0.94)**	0.92 (0.89-0.95)**	
Previous cesarean	43.91 (42.16-45.74)**	57.55 (55.26-59.93)**	57.21 (54.94-59.59)**	61.02 (58.57-63.57)**	60.58 (58.15-63.12)**	
Twins or multiple births	8.86 (8.16-9.61)**	12.09 (11.15-13.11)**	12.23 (11.29-13.26)**	12.32 (11.36-13.37)**	12.44 (11.47-13.50)**	
Gestational age						
< 37 w	1.87 (1.82-1.93)**	1.30 (1.26-1.35)**	1.33 (1.28-1.38)**	1.27 (1.22-1.32)**	1.29 (1.24-1.34)**	
≥ 37 w while < 42 w	(Reference group)	(Reference group)	(Reference group)	(Reference group)	(Reference group)	
≥ 42w	1.89 (1.82-1.94)**	1.97 (1.90-2.04)**	1.96 (1.88-2.03)**	2.07 (1.99-2.15)**	2.06 (1.98-2.14)**	
Fetal position						
Normal	(Reference group)	(Reference group)	(Reference group)	(Reference group)	(Reference group)	

(continued)

Table 2. Continued.

Variables/Logistic regressions	Multivariate		Hierarchical-multivariate		
	Univariate	Traditional	Streamlined	Traditional	Streamlined
Abnormal head position	6.73 (6.17–7.34)**	13.56 (12.42–14.80)**	13.56 (12.43–14.80)**	15.18 (13.89–16.59)**	15.14 (13.85–16.54)**
Breech or transverse position	13.28 (12.35–14.29)**	21.89 (20.37–23.52)**	21.73 (20.23–23.35)**	23.83 (22.16–25.61)**	23.65 (22.00–25.42)**
Cephalopelvic disproportion	93.04 (76.33–113.39)**	211.44 (174.02–256.91)**	210.43 (173.19–255.68)**	240.66 (197.99–292.53)**	239.14 (196.74–290.68)**
Umbilical cord entanglement	1.38 (1.12–1.16)**	1.41 (1.38–1.43)**	1.40 (1.38–1.43)**	1.40 (1.38–1.43)**	1.40 (1.37–1.43)**
Placenta previa	18.53 (16.84–20.40)**	17.57 (16.03–19.26)**	17.44 (15.91–19.12)**	18.06 (16.47–19.81)**	17.90 (16.33–19.64)**
Placental implantation and adhesion	1.96 (1.91–2.02)**	1.41 (1.36–1.46)**	1.42 (1.37–1.47)**	1.35 (1.30–1.40)**	1.35 (1.31–1.40)**
Placental abruption	3.14 (2.85–3.46)**	4.20 (3.78–4.67)**	4.22 (3.80–4.68)**	4.30 (3.86–4.78)**	4.31 (3.88–4.79)**
Premature rupture of membranes	0.60 (0.59–0.62)**	0.79 (0.77–0.80)**	0.78 (0.76–0.79)**	0.78 (0.76–0.80)**	0.77 (0.76–0.79)**
Situation of labor process					
Normal	(Reference group)	(Reference group)	(Reference group)	(Reference group)	(Reference group)
Precipitate labor	0.003 (0.002–0.006)**	0.002 (0.001–0.004)**	0.002 (0.001–0.005)**	0.002 (0.001–0.005)**	0.002 (0.001–0.005)**
Prolonged labor process	2.38 (2.10–2.70)**	4.18 (3.67–4.76)**	4.22 (3.70–4.80)**	4.26 (3.74–4.86)**	4.30 (3.77–4.91)**
Medical conditions level of pregnancy					
None	(Reference group)	(Reference group)	(Reference group)	(Reference group)	(Reference group)
Mild	1.13 (1.09–1.17)**	-	1.32 (1.26–1.38)**	-	1.29 (1.23–1.34)**
Moderate	2.27 (2.18–2.35)**	-	2.99 (2.87–3.12)**	-	3.04 (2.91–3.17)**
Severe	1.89 (1.85–1.92)**	-	2.02 (1.97–2.07)**	-	1.98 (1.94–2.03)**
Incompetent cervix	2.50 (2.35–2.66)**	2.42 (2.26–2.61)**	-	2.45 (2.28–2.64)**	-

(continued)

Table 2. Continued.

Variables/Logistic regressions	Multivariate			Hierarchical-multivariate		
	Univariate	Traditional	Streamlined	Traditional	Streamlined	Streamlined
Preeclampsia	2.09 (1.96-2.23)**	1.26 (1.17-1.36)**	-	1.32 (1.22-1.42)**	-	-
Any hypertension	2.34 (2.26-2.44)**	2.64 (2.53-2.76)**	-	2.68 (2.57-2.80)**	-	-
Diabetes mellitus	1.69 (1.65-1.72)**	1.61 (1.57-1.65)**	-	1.57 (1.53-1.61)**	-	-
Cardiac diseases	3.13 (2.77-3.54)**	4.32 (3.80-4.91)**	-	4.04 (3.55-4.59)**	-	-
Acute or chronic lung disease	1.18 (1.05-1.33)**	1.37 (1.18-1.58)**	-	1.30 (1.12-1.52)**	-	-
Hepatitis B	1.19 (1.07-1.31)**	1.14 (1.01-1.28)*	-	1.18 (1.05-1.34)*	-	-
Intrahepatic cholestasis	2.67 (2.51-2.84)**	3.61 (3.37-3.86)**	-	3.66 (3.42-3.92)**	-	-
Thyroid dysfunction	1.16 (1.12-1.20)**	1.26 (1.21-1.31)**	-	1.21 (1.16-1.26)**	-	-
Syphilis	1.53 (1.32-1.77)**	1.37 (1.15-1.63)**	-	1.39 (1.17-1.66)**	-	-

* $P < 0.05$; ** $P < 0.01$.

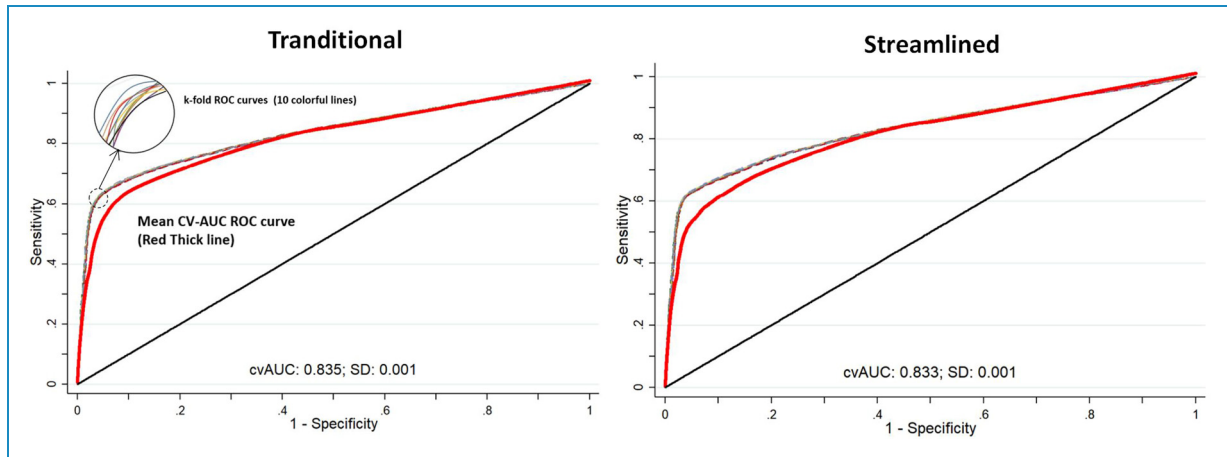


Figure 3. Comparison chart of two RCSR models in cross-validation.

Table 3. The evaluation of the effectiveness of two models.

Evaluation index/Type of model	Traditional	Streamlined
AUC	0.840	0.839
Accuracy	0.875	0.872
Sensitivity	0.690	0.685
Specificity	0.898	0.892
Positive predictive value	0.908	0.903
Negative predictive value	0.664	0.660
Brier score	0.069	0.067

determinants influencing CSR, during the process of institutional ranking aimed at measuring management quality, these maternal factors are precisely identified as variables requiring adjustment. This adjustment facilitates a more straightforward comparison of the actual quality (especially in management) of obstetric care among hospitals.

In addition, this authoritative study also accentuates the imperative for prospective research aimed at devising hospital-level interventions to refine the utilization of cesarean procedures.³⁸ The suggested protocol exhibits a substantial degree of alignment with the conceptual underpinnings of our scholarly inquiry. We have operationalized the antecedent components of the hospital-level intervention, specifically the evaluative processes, by meticulously enhancing a suite of algorithms designed for the risk stratification and adjustment of CSR. Then the subsequent obstetric medical quality intervention initiatives aimed at the rational control of cesarean delivery rates may well be predicated upon the assessment outcomes derived from this research.

Result of the two models. In addressing the prevalent complexity inherent within the traditional RCSR model, our study embarked on a systematic classification of medical conditions in pregnancy, adhering to obstetric clinical guidelines recognized at national and international levels. This structured approach underpinned the development of a streamlined RCSR model. It was subsequently subjected to comparative analysis against its traditional counterpart, evaluating statistical robustness and pragmatic efficacy within a hospital setting. The comparative outcomes indicated a substantial concordance between the two models regarding evaluative performance metrics and the resultant hospital rankings. Such parallelism provided a compelling rationale for the potential substitution of the traditional model with the streamlined model. The latter's adoption promises to expedite the assessment of obstetric care quality while preserving its integrity, thereby enhancing the overall expediency of medical quality evaluation processes.

Statistical methodology. Owing to the aggregation of datasets from 56 public tertiary healthcare institutions across Hubei Province, the present investigation qualifies as a multicentric study. An inherent challenge in this study design is the clustering phenomenon, where patient distributions across hospitals may be non-random due to a myriad of factors, including but not limited to the institutions' reputational standing, specialization, human resource deployment, and capital investment in medical equipment.³⁹ This dataset is acknowledged to possess hierarchical structural characteristics, necessitating analytical approaches adept at navigating such complexities. In addressing the nuances of hierarchical data structures, the employment of hierarchical modeling emerges as a methodologically sound approach.⁴⁰ A pivotal tool in quantifying the extent of clustering within the data is the intraclass correlation coefficient (ICC),^{41,42} which spans a continuum from 0 to

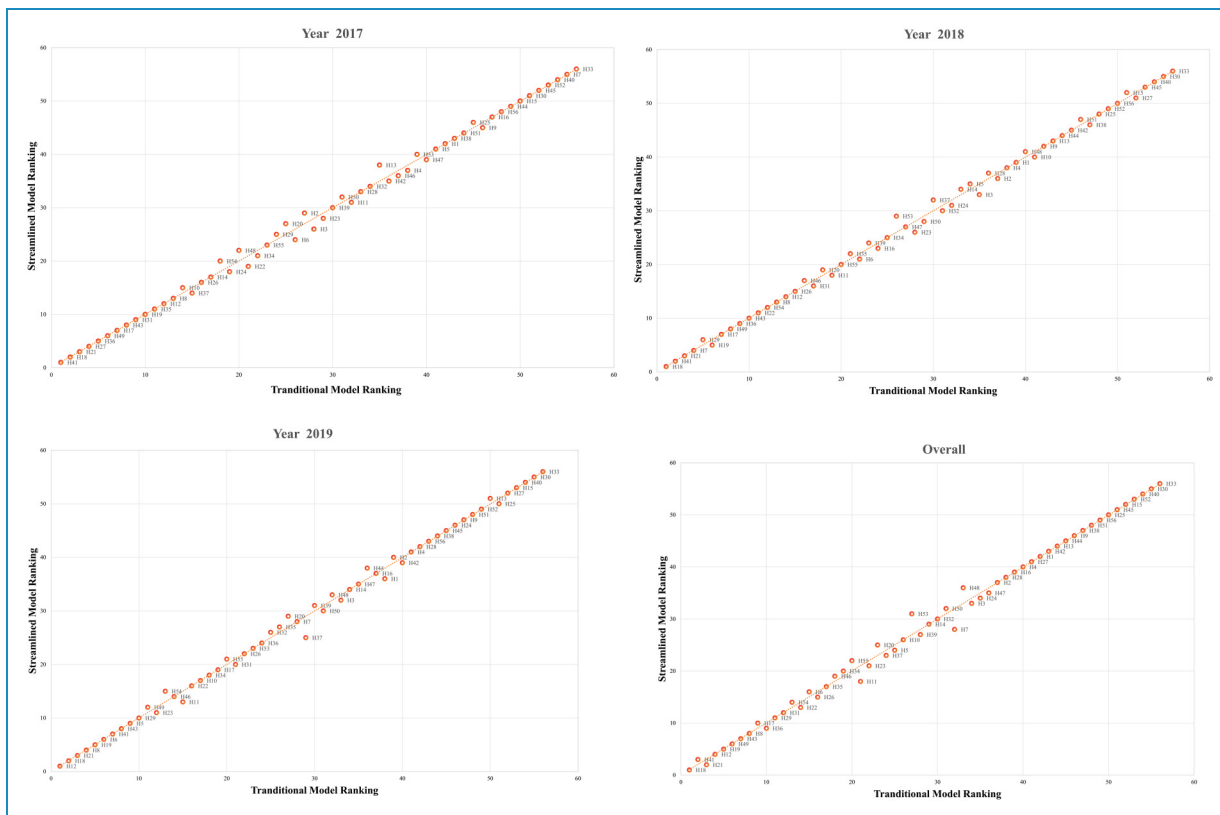


Figure 4. Annual changes in rankings under two models.

Table 4. The consistency test results of hospital annual rankings under two models.

Options/Years	2017	2018	2019	Overall
Kendall's tau coefficient	0.979**	0.978**	0.978**	0.974**
P-value	< 0.001	< 0.001	< 0.001	< 0.001

*P < 0.05; **P < 0.01.

1. There exists a scholarly consensus positing that an ICC value of 0.4 or greater signifies a notable clustering effect, warranting the incorporation of this dimensional complexity into the analytical model.^{43–45} Conversely, a corpus of research posits that neglecting to account for clustering, even when ICC values fall below 0.4, does not obviate the risk of incurring Type I errors in the resultant statistical analysis.^{46–48} Consequently, despite the ICC value being delineated at 0.18 in our study, significantly beneath the conventional threshold signaling pronounced clustering, our investigation meticulously adopted a hierarchical modeling framework. This deliberate methodological decision was predicated on the imperative to comprehensively ameliorate the conceivable confounding influences attributable to clustering

at the level of parturients, thereby safeguarding the veracity and methodological rigor of our analytical findings.

In the present analysis, the hierarchical model is synonymous with a mixed effects model, wherein factors associated with parturients are treated as fixed effects, while hospital-level variables (such as the degree of service refinement, geographic location, and teaching status) are conceptualized as random effects.⁴⁹ However, in the actual application process, apart from the codes of various hospitals, we did not incorporate the actual features at the hospital level into the model. This omission stems from the premise that, within the framework of risk adjustment, factors pertinent to the hospital level are deemed to exert an indirect (not direct) influence on concrete medical outcomes.¹² Another reason is that the primary objective of this research is to attenuate the effects of patient-level factors and clustering to the greatest extent feasible, rather than delineating the influence that relative hospital-specific factors may exert upon the actual outcomes of medical interventions.⁵⁰

As a modest methodological advancement, this investigation is underpinned by a hierarchical model framework and delves into the synergistic application of cross-validation and risk adjustment methodologies. Employing a hierarchical logistic regression model as the cornerstone

for modeling establishment, this approach entails the selection of an optimal training dataset, predicated upon the performance metrics of the corresponding validation fold. This selected dataset is then utilized forthwith in the construction of risk-adjusted prediction models. This holistic strategy adeptly navigates three pivotal challenges encountered during model development: the clustering of maternal data, enhancement of model optimization, and refinement of model training protocols. By adopting this innovative method, it is possible to circumvent computational discrepancies engendered by a multitude of factors, thereby facilitating the attainment of exemplary risk adjustment for quality metrics, predicated on binary medical outcomes, across multiple centers.

Digital quality evaluation based on EHRs

In the context of China's ongoing commitment to the high-quality development objectives of public hospitals, the initiative to construct intelligent hospital systems, including EMRs, has attained an unparalleled zenith.⁵¹ In 2018, the National Institute of Hospital Administration (NIHA), operating under the auspices of the National Health Commission of the People's Republic of China, launched a systematic program for evaluating and grading EMRs. This initiative, grounded in the guidelines articulated by the Society for Medical Information and Management Systems (HIMSS),⁵² is being methodically implemented across all hospital institutions at a national level on an annual basis.⁵³ This endeavor aims to standardize and enhance the quality and efficiency of EMR systems across the healthcare landscape.

As the level of Chinese EMR evaluation and grading advances, for hospitals in China, there is not only an augmentation in their functional diversity but, more paramount, a marked improvement in the caliber of data aggregation. The elevated data quality inherent in EMRs is an essential precondition for implemented digital medical quality evaluations. Consequently, the hospitals chosen for our study were solely those situated within Hubei Province that have attained a level 3 or higher grading evaluation for their EMRs. This selection criterion ensures the dependability and meticulousness of our evaluation and research efforts.

This research delineates a digital pathway for assessing obstetric medical quality utilizing EMRs, with a specific focus on cesarean sections as a pivotal metric. Additionally, this research not only undertakes merely the optimization of the conventional RCSR model but also significantly enhances the risk adjustment methodologies within the statistical analysis, particularly at the predictive modeling phase. Therefore, while this study inaugurates research within the domain of obstetrics, its implications extend far beyond. Crucially, it facilitates the prospect of conducting comprehensive, cross-regional, and multicenter

evaluations of medical quality through intelligent systems. This contribution offers invaluable empirical insights for health management entities. As part of a broader series of research endeavors, this study anticipates the integration of advanced algorithms, including artificial intelligence and machine learning, to further refine and expand the analytical framework presented herein.²⁵

Limitations

A notable limitation of this study is the paucity of data regarding maternal willingness, as EMRs do not adequately capture the psychological dimensions of postpartum women, absent a targeted investigation. Furthermore, even hospitals employing EMRs with a grading of level 3 or higher are not completely immune to challenges pertaining to data quality, underscoring the need for further refinement and standardization of EMR systems. Additionally, the inclusion of specialized obstetric hospitals introduces a potential source of bias. However, the primary focus of this research is to assess the feasibility of the proposed method. While such bias might be mitigated within the hierarchical model, it remains an aspect requiring acknowledgment.

Conclusion

In the realm of model performance evaluation as well as the pragmatic application within hospital settings, the streamlined model exhibits a substantial congruence with the traditional model. Therefore, the streamlined model can effectively serve as a viable surrogate for the traditional model, potentially establishing itself as a refined paradigm for the appraisal of quality in obstetric healthcare services.

Acknowledgements: The authors thank the Shenzhen Hospital Managers Association for their support of this study and the School of Medicine and Health Management of Tongji Medical College of Huazhong University of Science and Technology for their help in policy analysis. The authors extend their gratitude to Dr Huiwen Xu from the School of Public and Population Health, University of Texas Medical Branch, for his invaluable guidance in the construction of the research model.

Contributorship: DZ, CX, and JW led the design of the study and drafted the manuscript. JW, WZ, QS, and QF participated in the interpretation of results and drafting of the manuscript. DZ contributed to data acquisition. CX and BL conducted data acquisition and statistical analysis. SZ, CX, YS, QS, and JW provided professional obstetrical clinical consultation. PL and GL participated in the collation of big clinical data in this study. MW and XC revised and embellished the manuscript professionally. JW and CX serve as co-corresponding authors. DZ, WZ, and QS contributed equally to this paper as the co-first authors.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval: This study has been granted ethical approval by the Ethics Review Committee of Zhuhai People's Hospital (Zhuhai Hospital Affiliated with Jinan University), which includes a waiver of the requirement for informed consent (ethical approval number: LW-[2021]NO.7).

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Guangdong Provincial Health Economics Association, Peking University Shenzhen Hospital Management Research Project Fund, Shenzhen Municipal Health Economics Association (grant numbers 2022-WJMF-24, JCYJ2021009, NO.202265).

Guarantor: DZ.

ORCID iD: Chang Xu  <https://orcid.org/0000-0003-3323-9918>

Supplemental material: Supplemental material for this article is available online.

References

- Li H, Dong S, Liao Z, et al. Retrospective analysis of medical malpractice claims in tertiary hospitals of China: the view from patient safety. *BMJ Open* 2020; 10: e34681.
- Xu C, Zhong W, Fu Q, et al. Differential effects of different delivery methods on progression to severe postpartum hemorrhage between Chinese nulliparous and multiparous women: a retrospective cohort study. *BMC Pregnancy Childbirth* 2021; 21: 64.
- Xu C, Fu Q, Tao H B, et al. Effect of cesarean section on the severity of postpartum hemorrhage in Chinese women: the Shanxi study. *Curr Med Sci* 2018; 38: 618–625.
- Howell EA, Zeitlin J, Hebert PL, et al. Association between hospital-level obstetric quality indicators and maternal and neonatal morbidity. *JAMA* 2014; 312: 1531.
- Clark DE, Hannan EL and Wu C. Predicting risk-adjusted mortality for trauma patients: logistic versus multilevel logistic models. *J Am Coll Surg* 2010; 211: 224–231.
- Zhang J, Geerts C, Hukkelhoven C, et al. Caesarean section rates in subgroups of women and perinatal outcomes. *Br J Obstet Gynecol* 2016; 123: 754–761.
- Bailit JL, Garrett J M, Miller WC, et al. Hospital primary cesarean delivery rates and the risk of poor neonatal outcomes. *Am J Obstet Gynecol* 2002; 187: 721–727.
- Bailit JL, Dooley SL and Peaceman AN. Risk adjustment for interhospital comparison of primary cesarean rates. *Obstet Gynecol* 1999; 93: 1025–1030.
- Gibson K, Bailit JL, et al. Cesarean delivery as a marker for obstetric quality. *Clin Obstet Gynecol* 2015; 58: 211–216.
- Bailit JL and Garrett JM. Stability of risk-adjusted primary cesarean delivery rates over time. *Am J Obstet Gynecol* 2004; 190: 395–400.
- Bailit JL, Love TE and Dawson NV. Quality of obstetric care and risk-adjusted primary cesarean delivery rates. *Am J Obstet Gynecol* 2006; 194: 402–407.
- Bailit JL, Grobman WA, Rice MM, et al. Risk-adjusted models for adverse obstetric outcomes and variation in risk-adjusted outcomes across hospitals. *Am J Obstet Gynecol* 2013; 209: 441–446.
- Berghella V, Baxter JK and Chauhan SP. Evidence-based labor and delivery management. *Am J Obstet Gynecol* 2008; 199: 445–454.
- Nippita TA, Lee YY, Patterson JA, et al. Variation in hospital caesarean section rates and obstetric outcomes among nulliparae at term: a population-based cohort study. *BJOG: An Int J Obstet Gynaecol* 2015; 122: 702–711.
- Liu S, Liston RM, Joseph KS, et al. Maternal mortality and severe morbidity associated with low-risk planned cesarean delivery versus planned vaginal delivery at term. *CMAJ* 2007; 176: 455–460.
- Fullerton G, Danielian PJ and Bhattacharya S. Outcomes of pregnancy following postpartum haemorrhage. *BJOG* 2013; 120: 621–627.
- Bailit JL. Measuring the quality of inpatient obstetrical care. *Obstet Gynecol Surv* 2007; 62: 207–213.
- Birene B, Ishaque U, Chrusciel J, et al. Influence of the external cephalic version attempt on the cesarean section rate: experience of a type 3 maternity hospital in France. *Arch Gynecol Obstet* 2021; 303: 443–454.
- Quibel T, Chesnais M, Bouyer C, et al. Cesarean section rate changes after audit and feedback with the Ten Group Classification System in a French perinatal network: a retrospective pre-post study. *Acta Obstet Gynecol Scand* 2022; 101: 388–395.
- Schuster T, Lowe WK and Platt RW. Propensity score model overfitting led to inflated variance of estimated odds ratios. *J Clin Epidemiol* 2016; 80: 97–106.
- ACOG. ACOG Committee Opinion No. 447. Patient safety in obstetrics and gynecology. *Obstet Gynecol* 2009; 114: 1424–1427.
- China N H A F. Specification for risk assessment and management of parturient[EB/OL]. 2017. [\(http://www.nhc.gov.cn/fys/s3582/201711/f40145d5d1a84e998047c33c31ebd12d.shtml\)](http://www.nhc.gov.cn/fys/s3582/201711/f40145d5d1a84e998047c33c31ebd12d.shtml). (2017-11-17)[2024-02-15].
- Small DW, Sundstrom EJ and Head-Gordon M. A simple way to test for collinearity in spin symmetry broken wave functions: general theory and application to generalized Hartree Fock. *J Chem Phys* 2015; 142: 94112.
- Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. *JAMA* 2017; 318: 1377–1384.
- Xu H, Bowblis JR, Becerra AZ, et al. Developing a machine learning risk-adjustment method for hospitalizations and emergency department visits of nursing home residents with dementia. *Med Care* 2023; 61: 619–626.
- Paul P, Pennell ML and Lemeshow S. Standardizing the power of the Hosmer-Lemeshow goodness of fit test in large data sets. *Stat Med* 2013; 32: 67–80.
- Stata Statistical Software: Release 14[CP/OL].

28. Bailit JL, Gregory KD, Srinivas S, et al. Society for Maternal-Fetal Medicine (SMFM) special report: current approaches to measuring quality of care in obstetrics. *Am J Obstet Gynecol* 2016; 215: B8–B16.
29. Deneux-Tharoux C, Carmona E, Bouvier-Colle MH, et al. Postpartum maternal mortality and cesarean delivery. *Obstet Gynecol* 2006; 108–3: 541–548.
30. Gao Y, Xue Q, Chen G, et al. An analysis of the indications for cesarean section in a teaching hospital in China. *Eur J Obstet Gynecol Reprod Biol* 2013; 170: 414–418.
31. Anonymous. Appropriate technology for birth. *Lancet (London, England)* 1985; 2: 436–437.
32. Molina G, Weiser TG, Lipsitz SR, et al. Relationship between cesarean delivery rate and maternal and neonatal mortality. *JAMA* 2015; 314: 2263–2270.
33. Li HT, Luo S, Trasande L, et al. Geographic variations and temporal trends in cesarean delivery rates in China, 2008–2014. *JAMA* 2017; 317: 69–76.
34. Li HT, Hellerstein S, Zhou YB, et al. Trends in cesarean delivery rates in China, 2008–2018. *JAMA* 2020; 323: 89–91.
35. Qiao J, Wang Y, Li X, et al. A Lancet commission on 70 years of women’s reproductive, maternal, newborn, child, and adolescent health in China. *Lancet* 2021; 397: 2497–2536.
36. Lumbiganon P, Laopaiboon M, Gulmezoglu AM, et al. Method of delivery and pregnancy outcomes in Asia: the WHO global survey on maternal and perinatal health 2007–08. *Lancet* 2010; 375: 490–499.
37. Mi J and Liu F. Rate of caesarean section is alarming in China. *Lancet* 2014; 383: 1463–1464.
38. Yin S, Chen L, Zhou Y, et al. Evaluation of cesarean rates for term, singleton, live vertex deliveries in China in 2020 among women with no prior cesarean delivery. *JAMA Netw Open* 2023; 6: e234521.
39. Griffiths P, Maruotti A, Recio SA, et al. Nurse staffing, nursing assistants and hospital mortality: retrospective longitudinal cohort study. *BMJ Qual Saf* 2019; 28: 609–617.
40. Zewude BT, Debusho LK and Diriba TA. Multilevel logistic regression modelling to quantify variation in malaria prevalence in Ethiopia. *PLoS One* 2022; 17: e273147.
41. Martinez PJ and Perez MP. [Intraclass correlation coefficient]. *Semergen* 2023; 49: 101907.
42. Shieh G. Choosing the best index for the average score intraclass correlation coefficient. *Behav Res Methods* 2016; 48: 994–1003.
43. Gulliford MC, Adams G, Ukoumunne OC, et al. Intraclass correlation coefficient and outcome prevalence are associated in clustered binary data. *J Clin Epidemiol* 2005; 58: 246–251.
44. Srivastava MS. Estimation of the intraclass correlation coefficient. *Ann Hum Genet* 1993; 57: 159–165.
45. Shieh G. A comparison of two indices for the intraclass correlation coefficient. *Behav Res Methods* 2012; 44: 1212–1223.
46. Hox JJ. *Multilevel Analysis: Techniques and Applications*. London: Psychology Press, 2002.
47. Leyrat C, Morgan KE, Leurent B, et al. Cluster randomized trials with a small number of clusters: which analyses should be used? *Int J Epidemiol* 2018; 47: 321–331.
48. Kovacic J and Varnai VM. Intraclass correlation coefficient for grouped data. *Epidemiology* 2014; 25: 769–770.
49. Tesema GA, Yeshaw Y, Kasie A, et al. Spatial clusters distribution and modelling of health care autonomy among reproductive-age women in Ethiopia: spatial and mixed-effect logistic regression analysis. *BMC Health Serv Res* 2021; 21: 74.
50. Massarweh NN, Chen VW, Rosen T, et al. Comparative effectiveness of risk-adjusted cumulative sum and periodic evaluation for monitoring hospital perioperative mortality. *Med Care* 2021; 59: 639–645.
51. Mi D, Li Y, Zhang K, et al. Exploring intelligent hospital management mode based on artificial intelligence. *Front Public Health* 2023; 11: 1182329.
52. Sun N, Li H, Jin Y, et al. Research on the application of HIMSS EMRAM 7-level multicode management. *Minerva Surg* 2023; 78: 601–604.
53. China T N H C. Notice on Issuing the Management Measures for Grading and Evaluating the Application Level of Electronic Medical Record Systems (Trial) and Evaluation Standards (Trial)[EB/OL].(2018-12-30)[2024-03-23]. <http://www.nhc.gov.cn/cms-search/xxgk/getManuscriptXxgk.htm?id=3cae6834a65d48e9bfd783f3c7d54745>.