A Clinical Prediction Rule for the Severity of Congenital Diaphragmatic Hernias in Newborns

**AUTHORS:** Mary Elizabeth Brindle, MD, MPH, Earl Francis Cook, ScD, Dick Tibboel, MD, PhD, Pamela A. Lally, MD, and Kevin P. Lally, MD, MSc, on behalf of the Congenital Diaphragmatic Hernia Study Group

**ABBREVIATIONS**

CDH—congenital diaphragmatic hernia

**WHAT’S KNOWN ON THIS SUBJECT:** Predicting high-risk populations in congenital diaphragmatic hernia (CDH) can help target care strategies. Prediction rules for infants with CDH often lack validation, are aimed at a prenatal population, and are of limited generalizability. We cannot currently discriminate the highest risk neonates during the crucial period shortly after birth.

**WHAT THIS STUDY ADDS:** This clinical prediction rule was developed and validated on an international database. It discriminates patients and high, intermediate, and low risk of mortality, is easy to apply, and is generalizable to most infants with CDH.

**BACKGROUND:** Congenital diaphragmatic hernia (CDH) is a condition with a highly variable outcome. Some infants have a relatively mild disease process, whereas others have significant pulmonary hypoplasia and hypertension. Identifying high-risk infants postnatally may allow for targeted therapy.

**METHODS:** Data were obtained on 2202 infants from the Congenital Diaphragmatic Hernia Study Group database from January 2007 to October 2011. Using binary baseline predictors generated from birth weight, 5-minute Apgar score, congenital heart anomalies, and chromosome anomalies, as well as echocardiographic evidence of pulmonary hypertension, a clinical prediction rule was developed on a randomly selected subset of the data by using a backward selection algorithm. An integer-based clinical prediction rule was created. The performance of the model was validated by using the remaining data in terms of calibration and discrimination.

**RESULTS:** The final model included the following predictors: very low birth weight, absent or low 5-minute Apgar score, presence of chromosomal or major cardiac anomaly, and suprasystemic pulmonary hypertension. This model discriminated between a population at high risk of death (~50%), intermediate risk (~20%), or low risk (<10%). The model performed well, with a C statistic of 0.806 in the derivation set and 0.769 in the validation set and good calibration (Hosmer-Lemeshow test, P = .2).

**CONCLUSIONS:** A simple, generalizable scoring system was developed for CDH that can be calculated rapidly at the bedside. Using this model, intermediate- and high-risk infants could be selected for transfer to high-volume centers while infants at highest risk could be considered for advanced medical therapies. *Pediatrics* 2014;134:e413–e419
Congenital diaphragmatic hernia (CDH) is a condition that is associated with significant morbidity and mortality as well as a proportionately high cost to the health care system. The overall survival of infants with this condition is highly variable as is their care, although the standardization of postnatal care has evolved over the years. This is illustrated by the dramatic differences in survival between infants undergoing patch repair compared with those amenable to primary repair. A method to identify these high- and low-risk infants in the postnatal period but early in the course of their disease would offer the opportunity to tailor therapies based on severity. Unfortunately, such a model has been elusive. A well-constructed clinical predictive rule can stratify these infants into populations at high and low risk of death. This method would aid in the identification of infants who might benefit from transfer to high-volume centers. In addition, identification of those infants at the greatest risk of mortality will help select an appropriate target population for trials and targeted therapies.

Prediction rules have previously been developed to predict outcomes for infants with CDH within the Congenital Diaphragmatic Hernia Study Group (CDHSG) database and others. Many of these models have been developed in small populations, are difficult to apply, and suffer from a lack of validation, poor discrimination, or poor generalizability. The aim of the present study was to create and validate a generalizable and easily applied clinical prediction rule to determine infants at high, intermediate, and low risk of death by using simple clinical parameters as well as echocardiographic measures of pulmonary hypertension.

Methods
Data were acquired from the CDHSG, and institutional review board approval was attained (University of Calgary institutional review board, ethics no. 20776). The CDHSG data set is an international, disease-specific registry including 59 centers across 10 countries goal is to include all neonates with the diagnosis of CDH from the time of diagnosis to discharge. Data are entered manually and cross-checked for accuracy. Missing data points are queried back to the center for availability; inconsistent data points are queried for accuracy. Data are reported back to the center periodically for center quality assurance and validation. The total data set comprises >7000 subjects. Beginning in January 2007, data were collected regarding echocardiographic measures of pulmonary hypertension.

We reviewed the data on patients entered into the database between January 2007 and October 2011. Data fields include those variables that would be available at the time of first echocardiogram. These consist of demographic information and known neonatal predictors of mortality, including gestational age, birth weight, Apgar score, inborn/outborn status, major cardiac anomalies (classified as all anomalies other than patent foramen ovale or patent ductus arteriosus), chromosomal anomalies, and gender; also included were features pertaining to the diaphragmatic defect such as side of defect and prenatal liver position. In addition, measures of pulmonary hypertension estimated on the first echocardiogram were included (echocardiogram was performed at a mean of 0.7 day of life in the data set). Specifically, presence of right to left shunting through a patent ductus arteriosus or estimates of pulmonary pressures were used. Echocardiographic measures of pulmonary hypertension were recoded to create composite measures, combining similar predictors as a binary variable. We created a binary variable of high pulmonary pressure in which this pressure was defined as right to left shunting or pulmonary pressures estimated as higher than systemic pressure reflecting previously described echocardiographic definitions of pulmonary hypertension.

Potential predictors were recoded into binary variables based on appropriate clinical cutoffs established a priori or by the median measure of the variable to create a clinical prediction rule that is easily applied. A total of 12 potential predictor variables were generated. The data set was divided into 2 groups by using a computerized 1-to-1 randomization process. One group was designated as a derivation set for developing the prediction model and the other was a testing or validating set. The crude distributions of baseline predictors for those who did and did not die were examined through univariate analysis. These were not used to determine entrance into the clinical predictive model, which was generated by using all potential predictors.

Variables that had >5% of data points missing were recreated as indicator variables comprising a value for “missing” and included within logistic regression analysis performed on the derivation set. Selection of the manner in which variables would be included in the analysis was made depending on the predictiveness of “missingness” within the model and how that could be interpreted in a clinical context.

A logistic regression with a backward elimination algorithm was performed to create a prediction model with a significance level of 0.01 for exclusion. Presence or absence of severe pulmonary hypertension was forced into the model. This model was developed on the derivation population and tested on the validation population. The performance of this model in identifying those patients at high risk of death was tested by comparing the frequency of death.
before discharge in the population believed to be at high risk (>50% risk of mortality) as well as those estimated at very high risk of death according to the prediction model (>75% risk) in both derivation and validation models. The calibration of the model was further assessed by using the Hosmer-Lemeshow goodness-of-fit test. Discrimination of the model was evaluated through assessment of the area under the receiver operating curve (C statistic). Further assessment of the degree of optimism of the model was obtained through cross-validation. The model was restricted to those infants having data on echocardiographic measures of pulmonary hypertension. Including an outcome measure of “missing” for all variables during model creation assessed the impact of missing data.

Based on the significant coefficients in the clinical predictive model, an equation was created dividing all coefficients by the lowest significant coefficient and rounding to the closest whole integer. These were used to provide a “score” for individual patients within the derivation group. Based on the range of scores and mortality rates for these infants, low-, intermediate-, and high-risk categories were created based on the total score, and the ability for this score to predict outcome was evaluated in the validation group.

RESULTS

Infants in the CDHSG database have a mortality rate of 28%. They are generally term infants with a median gestational age of 38 weeks and a median birth weight of 3.00 kg. Associated cardiac anomalies are uncommon (28%), and major cardiac anomalies and chromosomal anomalies are rare (8.3% and 4.6%, respectively).

With the exception of the side of the CDH defect, baseline characteristics of infants with CDH who died were significantly different from those who did not die (Table 1). Neonatal factors associated with increased mortality on crude univariate analysis included low birth weight, prematurity, low Apgar score at 5 minutes, presence of major congenital heart disease, and chromosomal abnormalities. In addition, the prenatal location of liver in the chest and increased measures of pulmonary hypertension were all associated with increased mortality in crude univariate analyses.

Logistic regression was performed within the derivation data set, including all variables with those having significant missing data transformed into indicator variables with a value for missingness. This analysis identified that missing data were predictive of mortality for 5-minute Apgar score, prenatal liver position, and echocardiographic findings. Missingness for gestational age was not associated with mortality. The liver position variable was removed from the analysis entirely, and the analysis was restricted to those patients with an echocardiogram performed and those with gestational age recorded. Five-minute Apgar score was recoded as a nominal variable with 3 measures (high, low, or missing) and included in the analysis.

The backward elimination algorithm with a significance of 0.01 for exclusion identified 6 predictors of mortality: low birth weight, low measured 5-minute Apgar score, the lack of a 5-minute Apgar score, the presence of major cardiac anomalies, the presence of chromosomal abnormalities, and the presence of significant pulmonary hypertension (Table 2). This prediction model demonstrated good discrimination based on the area under the receiver operating curve (C statistic, 0.806 in the derivation model and 0.769 in the validation model) (Table 3). The model was also found to be good at discriminating between those population at low, intermediate, and high risk of death (Table 4).

The model demonstrated good calibration as indicated by the Hosmer-Lemeshow goodness-of-fit test ($P = .217$) (Table 3). Cross-validation provided an estimation of optimism of the model of 0.065 for the C statistic, which is similar to the 0.037 found in the difference between the derivation and validation model.

The CDH scoring equation was created as described earlier by using all predictors with significant coefficients. All coefficients were then divided by 2.634 (the coefficient of low birth weight) to obtain the following equation:

$$1(\text{low birth weight}) + 1.051(\text{low Apgar}) + 1.77(\text{missing Apgar}) + 1.55(\text{severe pulmonary hypertension}) + 1.98(\text{major cardiac anomaly}) + 1.49(\text{chromosomal anomaly})$$

By rounding to the nearest integer, the following final equation was created, with a possible range of 0 to 8:

$$1(\text{low birth weight}) + 1(\text{low Apgar}) + 2(\text{missing Apgar}) + 2(\text{severe pulmonary hypertension}) + 2(\text{major cardiac anomaly}) + 1(\text{chromosomal anomaly}) = \text{total CDH score}$$

This score resulted in identification of infants at low, intermediate, and high risk of death (<10%, 25%, or 50%) based on the total score (0, 1–2, or ≥3). The distribution of scores was decided based on their performance in the derivation group (Table 5). The discrimination of this simplified model still performed well, as similar mortality predictions were found within the validation group (Table 4). The C statistic demonstrated fair discrimination of the points-based model with a value of 0.721.
TABLE 1 Baseline Characteristics of Infants With CDH Who Did and Did Not Die (N = 2022)

| Baseline Characteristics | All Patients (N = 2022) | No. Missing Data | Patients Who Died (n = 561) | Patients Who Did Not Die (n = 1456) | P Value

| Male | 1264 (62.7) | 10 | 527 (58.5) | 933 (64.2) | <.0001 |
| Prenatal diagnosis present | 1343 (66.8) | 15 | 450 (89.5) | 888 (61.4) | <.0001 |
| Liver in the chest on prenatal US | 709 (44.3) | 421 | 170 (77.3) | 539 (39.0) | <.0001 |
| Outborn | 1101 (54.5) | 6 | 262 (46.7) | 839 (57.7) | <.0001 |
| Major cardiac abnormality present | 168 (8.3) | 5 | 102 (18.2) | 66 (4.5) | <.0001 |
| Chromosomal abnormality present | 94 (4.7) | 5 | 60 (10.7) | 34 (2.3) | <.0001 |
| Other anatomic abnormality present | 351 (17.4) | 5 | 143 (25.5) | 208 (14.3) | <.0001 |
| Median gestational age at birth, wk | 38 (37–39) | 208 | 38 (35–39) | 38 (37–39) | <.0001 |

Data are presented as n (%) or median (interquartile range). GA, gestational age; US, ultrasound.

* For all binary comparisons, Fisher’s exact test was performed; for continuous variables, the Wilcoxon rank sum tests were applied.

b Right to left shunt or estimate of suprasystemic pulmonary pressures on echocardiogram.

DISCUSSION

We developed a validated clinical prediction rule that identifies infants at low, intermediate, and high risk of death at the time of the first echocardiogram. Given the international scope of this database and the range of participating institutions, this model will likely perform well in most infants with CDH.

There is considerable value in a prediction rule that can help identify populations at high risk for mortality. Prenatal imaging holds promise for prediction in many studies but is inconsistently available, and results suffer from poor reliability across centers.15–18 There are many postnatal studies that look at the factors associated with mortality in CDH, but there are few clinical prediction models for CDH.

A prediction model was developed from the CDHSG in 2001 containing the 5-minute Apgar score and birth weight.15 Although this model performed well when validated by using a separate data set,11 the performance of this original model is inferior to the current model in its discriminatory abilities and is somewhat more cumbersome to apply despite having only 2 variables. Schultz et al10 developed a clinical prediction model based on blood gas analysis. This model discriminated between a population with a 66% survival and a population with an 83% survival. Although the model performed well in the initial validation study, it did not perform well in further validation studies.11 The rule we have developed has been validated, and it performs well in terms of discrimination and calibration. It identified infants within different risk strata and is applied by using a points-based system.

There are a number of statistical elements that should be recognized when constructing and interpreting a clinical prediction rule. Variables in a clinical prediction rule do not need to be and are frequently not independent. There may also be interaction between variables that are not apparent in the application of the rule. These factors do not impact the performance of the constructed model but are important to recognize if the results were to be further extrapolated.

Generalizability of predictive equations is often greatly influenced by available data. Many infants do not have a 5-minute Apgar score within the CDHSG.
We were particularly interested in investigating the predictive value of echocardiographic data, and 292 (14.4%) patients were missing data on this variable. We have limited the application of this rule to infants who have undergone echocardiographic evaluation. It is likely that those infants missing data on echocardiographic estimates of pulmonary hypertension are systematically different from those with echocardiographic data. By restricting the applicability of the clinical prediction rule to those with echocardiographic data, we have likely eliminated a population that is at higher risk of mortality. This suggestion is supported by the subset analysis, which revealed that those patients missing echocardiographic data had a mortality rate of 38.7%.

There are a number of limitations to this study. We created this clinical prediction rule from the CDHSG database, which is the largest CDH-specific database in existence. However, the CDHSG is a voluntary database rather than a registry, and it therefore may not reflect the true population of infants with CDH. In addition, in this registry, there is no standardized method of reporting variables such as echocardiographic measures of pulmonary hypertension, and thus there is likely significant variability between centers. For simplicity, we created binary variables for our predictors, which produce a model that is easy to apply but loses some of its accuracy. One binary predictor created is that of pulmonary hypertension. This variable includes the presence of ductal shunting or relies on a reported estimation of pulmonary pressures without a standardized method of measurement prescribed. This method may lead to inaccuracies and variability but including right to left shunting through the patent ductus arteriosus as well as estimations of suprasystemic pressure is likely to reflect the population of infants with the highest pulmonary pressures. Because the echocardiographic measures were frequently attained in the first 24 hours, it could be anticipated that the residual effects of fetal circulation might diminish the significance of early echocardiogram findings. Although we did find echocardiographic measures to be predictive in our model, an echocardiogram performed at a later time frame may have a greater predictive value. We have included “missingness” of the 5-minute Apgar score as a predictive variable within our rule. Although this factor has proven to be a robust predictor within our validation set, we assumed that the reasons for missing an Apgar score would be similar for patients on whom we will apply this model, which may not always hold true.

Prenatal predictors offer great promise in terms of determining prognosis. In developing our model, it was clear that prenatal liver position was highly predictive of mortality but was very poorly reported within the CDHSG data set.
Patients missing these data also had higher mortality. Clinical prediction rules integrating prenatal predictors have not currently been developed, although many studies have been performed that demonstrate the value of prenatal predictors such as lung to head ratio, MRI estimations of lung volumes, and liver position. With advancements in prenatal imaging, a prenatal scoring system could help direct clinical care pathways for the fetus as well as the neonate. However, more consistent application and reporting of prenatal imaging would be required to make such a model generalizable. In the interim, it should be left to the judgment of the clinician on how to incorporate findings such as prenatal liver position depending on institutional practice and the reliability of that measure.

The variation in care across institutions is a limitation of the CDHSG data set. However, a prediction rule generated from this data set has greater generalizability because it reflects the current, variable population on whom it will be applied.

We have developed a clinical predictive rule that allows for identification of a large population of infants that can be successfully managed at low-volume centers closer to their families. In the validation population of 886, almost one-third (n = 287) were identified in a low-risk group with a mortality rate of <10%. In countries with a large geographic area, smaller centers will frequently be the first hospitals that manage infants with CDH. Increasingly, there is a recognition that centers with larger volumes may have improved outcomes in CDH. The burden on the family and health care system that would be involved in the transfer of all patients with CDH to high-volume centers could be diminished through identification of a population with a lower risk of mortality that can be cared for in smaller volume institutions; those at higher risk could then be considered for transport to those centers with greater expertise. This practice is already accepted in many European countries.

This rule also identifies a very high-risk population that can be targeted in practice guidelines and clinical trials. These are the patients for whom the greatest benefits are likely to be achieved in directed therapies. In small-volume clinical trials, methods of targeting studies to those patients at highest risk increases the likelihood that a true benefit will be detected as well as diminishing the cost and unnecessary exposure of a low-risk population to experimental therapies.

Standardized timing and methods of performing echocardiography may significantly improve the role of this powerful diagnostic test in predicting patient risk. Further validation of this model within an external population will be of benefit in future studies.

ACKNOWLEDGMENTS

The authors acknowledge Dr Shimon Sheykevich, who provided statistical programs used in the analyses, and the many centers that participate in the CDHSG.

REFERENCES


A Clinical Prediction Rule for the Severity of Congenital Diaphragmatic Hernias in Newborns
Mary Elizabeth Brindle, Earl Francis Cook, Dick Tibboel, Pamela A. Lally and Kevin P. Lally
Pediatrics 2014;134:e413
DOI: 10.1542/peds.2013-3367 originally published online July 14, 2014;
A Clinical Prediction Rule for the Severity of Congenital Diaphragmatic Hernias in Newborns
Mary Elizabeth Brindle, Earl Francis Cook, Dick Tibboel, Pamela A. Lally and Kevin P. Lally
Pediatrics 2014;134;e413
DOI: 10.1542/peds.2013-3367 originally published online July 14, 2014;

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://pediatrics.aappublications.org/content/134/2/e413