Journal of Pediatric Surgery (2008) 43, 783–787

ELSEVIER

Journal of Pediatric Surgery

www.elsevier.com/locate/jpedsurg

Mortality prediction in congenital diaphragmatic hernia

Robert Baird[a,b], Ying C. MacNab[c], Erik D. Skarsgard[a,b,*]
the Canadian Pediatric Surgery Network

[a]*Division of Pediatric Surgery, British Columbia Children's Hospital, Vancouver, British Columbia, Canada V6H 3V4*
[b]*Department of Surgery, University of British Columbia, Vancouver, British Columbia, Canada V5Z 4E3*
[c]*Department of Health Care and Epidemiology, University of British Columbia, Vancouver, British Columbia, Canada V6H 3V4*

Received 12 November 2007; accepted 3 December 2007

Key words:
Congenital diaphragmatic hernia;
Survival;
Risk-adjustment;
Predictive equation

Abstract

Background: A validated risk stratification tool for congenital diaphragmatic hernia (CDH) is required for accurate outcomes analyses. Existing mortality-predictive models include those of the CDH Study Group (CDHSG) based on birth weight and 5-minute Apgar score, the Canadian Neonatal Network (CNN) based on gestational age and admission score in Score for Neonatal Acute Physiology version II, and the Wilford Hall/Santa Rosa clinical prediction formula ($WHSR_{PF}$) derived from blood gas measurements. The purpose of this study was to evaluate the calibration and discrimination of these predictive models using the Canadian Pediatric Surgical Network dataset.

Methods: Neonatal risk variables and birth hospital survivorship were collected prospectively in 11 perinatal centers, between May 2005 and October 2006. Actual vs predicted outcomes were analyzed for each equation to measure the calibration and discrimination of each model.

Results: Twenty (21.2%) of 94 infants with CDH died during birth hospitalization. The CDHSG model demonstrated superior discrimination (area under the receiver operator characteristic curve = 0.85; CNN = 0.79; $WHSR_{PF}$ = 0.63). Model calibration reflected by the Hosmer-Lemeshow *P* value was poorest with the $WHSR_{PF}$ = 0.37 and comparable between CDHSG and CNN (0.48 and 0.46, respectively).

Conclusion: Predictive outcome models are essential for risk-adjusted outcome analysis of CDH. The ideal predictive equation should prove robust across CDH datasets.

Despite significant advances in prenatal diagnosis and neonatal intensive care, congenital diaphragmatic hernia (CDH), continues to be a vexing congenital malformation with broadly variable cardiopulmonary disease severity at birth. Although implementation of rational treatment strategies including preoperative stabilization, lung protective ventilation, extracorporeal membrane oxygenation (ECMO), pulmonary vasodilator therapy, and delayed surgical treatment have resulted in reduced mortality trends in individual centers, there continues to be great variability in overall CDH mortality with rates ranging between 20% and 60% [1-4].

An ongoing barrier to outcome research in CDH is the lack of a validated and widely accepted disease-severity adjustment tool. Validated illness-severity assessment of the newborn with CDH should be done early in the postnatal

Presented at the 39th Annual Meeting of the Canadian Association of Pediatric Surgeons, August 23-26, 2007, St John's Newfoundland, Canada.

* Corresponding author. Division of Pediatric General Surgery, KO-123 ACB, 448 Oak Street, Vancouver British Columbia, Canada V6H 3V4. Tel.: +1 604 875 3744; fax: +1 604 875 2721.
E-mail address: eskarsgard@cw.bc.ca (E.D. Skarsgard).

0022-3468/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jpedsurg.2007.12.012

course so that the outcome prediction provided is not significantly influenced by postnatal treatment. Early outcome prediction enables anticipation of the likelihood of the need for more aggressive treatment strategies (such as ECMO) and informs physician-parent discussions on an individual patient basis. From the perspective of studying populations of CDH infants, such a tool allows comparisons of risk-adjusted patients within and between institutions. Finally, accurate risk-adjusted outcome stratification enables the rigorous evaluation of treatment strategies in prospective trials.

There have been at least 3 predictive equations derived for early CDH mortality. These include the predictive equations of the CDH Study Group (CDHSG) [5], the Canadian Neonatal Network (CNN) [6], and the Wilford Hall/Santa Rosa group ($WHSR_{PF}$) [7]. Each predictive equation was derived from a unique CDH patient cohort. The purpose of our study was to validate each of these outcome predictors using a single CDH dataset from the Canadian Pediatric Surgical Network (CAPSNET) database.

1. Materials and methods

The dataset was obtained from the CAPSNET, a multidisciplinary group of 16 Canadian perinatal centers that collect prospective, disease-specific data on both CDH, and gastroschisis (Appendix 1). A perinatal center is defined as one with a level III neonatal intensive care unit (NICU), pediatric anesthesia, and subspecialty surgery (at least general and neurosurgery) capabilities, and a geographically or functionally adjoined maternal-fetal medicine/advanced prenatal diagnosis center. Pediatric cardiac surgery and ECMO are available in 9 and 4 CAPSNET centers, respectively. In Canada, all perinatal care for birth defects such as CDH is provided exclusively through these provincial referral centers.

Eleven centers contributed data during the study period; the largest site contributed 25 patients whereas 3 sites contributed 2 patients each.

1.1. Study population

Congenital diaphragmatic hernia cases for this study were accrued between May 1, 2005, and December 31, 2006. Cases were ascertained at prenatal diagnosis (if one was made) or after birth, and data were abstracted from diagnosis to death or discharge from a CAPSNET center. Infants transferred from one CAPSNET center to another were tracked back to their initial admission.

1.2. Data collection

Notification of the prenatal diagnosis or birth of a case of CDH was forwarded to on-site, trained research assistants who abstracted data from maternal and infant charts using a customized data entry program with built-in error checking and a standard manual of operations and definitions. Data were deidentifed and transmitted electronically to a centralized repository where data were cleaned, stored, and thereafter managed by a study coordinator and a geographically representative, multidisciplinary steering committee comprised of pediatric surgeons, neonatologists, maternal-fetal medicine specialists, and an epidemiologist.

1.3. Predictive models

Three predictive models were evaluated. Each model was derived from a separate CDH patient cohort, by testing (individually and in combination), those risk variables deemed to be predictive of mortality by multivariable logistic regression analysis within that cohort:

(1) The CDHSG probability of survival equation = $1-1/(1+e^{-x})$, where $-x = -5.0240 + 0.9165$ (birth weight in kilograms) + 0.4512 (Apgar score at 5 minutes) [5].

(2) The CNN predictive equation that uses a combination of 2 risk variables, the Score for Neonatal Acute Physiology version II (SNAP-II), and gestational age (GA). The SNAP-II is a standardized index, validated in other neonatal patient populations, which depicts illness severity by the magnitude of derangement in 6 physiologic parameters: mean blood pressure, lowest temperature, P_{O_2} (mm Hg)/F_{IO_2} (%) ratio, lowest serum pH, presence of seizure activity, and urine output (mL/kg per hour), expressed as an aggregate score [8].

(3) $WHSR_{PF}$. This equation uses blood gas values (from a primarily postductal source), measured during the first 24 hours of life to calculate the equation: highest Pa_{O_2}−highest P_{CO_2}, with a cutoff value of zero or greater expected to predict survival [7].

1.4. Data analysis

Perinatal characteristics, characteristics of operated patients, and outcomes were recorded directly from the CAPSNET database. For each predictive equation, modeled and actual outcomes were compared using the receiver operator characteristic (ROC) curve technique of Hanley and MacNeil [9] to assess model discrimination. The area under the ROC curve (AUC) depicts model discrimination—a measure of its predictive performance. An AUC of 0.5 represents a completely random association between the modeled and actual outcomes, whereas an AUC of 1.0 represents perfect discrimination. The conformity between actual and predicted outcome is also depicted by model calibration or "goodness of fit," where, using the Hosmer-Lemeshow (H-L) technique, a P value of .05 or higher suggests that there is no difference between modeled and actual outcomes. The higher the P value, the better the model calibration [10]. All analyses were performed using SPSS for Windows statistical software (SPSS, Chicago, Ill).