Journal of Pediatric Surgery 55 (2020) 2064–2070

Contents lists available at ScienceDirect

Journal of Pediatric Surgery

journal homepage: www.elsevier.com/locate/jpedsurg

Development and validation of risk models for mortality and morbidity in 12 major pediatric surgical procedures: A study from the National Clinical Database-Pediatric of Japan

Keita Terui [a,k,*], Norimichi Hirahara [b,k], Hisateru Tachimori [b,c,d], Naohiro Kato [c], Jun Fujishiro [e,k], Eiichiro Watanabe [e,k], Hirofumi Tomita [f,k], Tatsuya Okamoto [g,k], Michimasa Fujiogi [e,k], Shinya Okamoto [h,k], Takeo Yonekura [i,k], Hiroaki Miyata [b], Noriaki Usui [j,k]

[a] *Department of Pediatric Surgery, Graduate School of Medicine, Chiba University, Chiba, Japan*
[b] *Department of Healthcare Quality Assessment, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan*
[c] *Translational Medical Center, National Center of Neurology and Psychiatry, Tokyo, Japan*
[d] *The Institute for Global Health Policy Research, Bureau of International Health Cooperation, National Center for Global Health and Medicine, Tokyo, Japan*
[e] *Department of Pediatric Surgery, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan*
[f] *Department of Surgery, Tokyo Metropolitan Children's Medical Center, Tokyo, Japan*
[g] *Department of Hepato-Biliary-Pancreatic Surgery and Transplantation, Graduate School of Medicine, Kyoto University, Kyoto, Japan*
[h] *Department of Pediatric Surgery, Hyogo Prefectural Amagasaki General Medical Center, Amagasaki, Japan*
[i] *Department of Pediatric Surgery, Kindai University School of Medicine Nara Hospital, Nara, Japan*
[j] *Department of Pediatric Surgery, Osaka Women's and Children's Hospital, Izumi, Japan*
[k] *National Clinical Database Committee and [l] Pediatric Surgical Database Committee, The Japanese Society of Pediatric Surgeons, Tokyo, Japan*

ARTICLE INFO

Article history:
Received 19 September 2019
Received in revised form 22 February 2020
Accepted 27 March 2020

Key words:
Surgical procedures
Health care evaluation mechanisms
Postoperative complications
Intraoperative Complications
Mortality
Morbidity

ABSTRACT

Purpose: To establish and validate risk models of mortality and morbidity associated with 12 major pediatric surgical procedures using the National Clinical Database-Pediatric (NCD-P) data.
Methods: We used the NCD-P data for the development and validation datasets. By using multivariate logistic regression to analyze the development dataset, we created a prediction model for 30-day mortality and morbidity in 12 major pediatric surgical procedures, including tracheoplasty, pneumonectomy, fundoplication, total/subtotal excision of malignant tumor, and surgeries for Hirschsprung disease, anorectal malformation, biliary atresia, choledocal cyst, midgut volvulus, funnel chest, gastrointestinal perforation, and intestinal obstruction. We selected variables that were almost identical to those used in the American College of Surgeons National Surgical Quality Improvement Program-Pediatric (NSQIP-P). The primary outcomes were 30-day mortality and composite morbidity. We assessed the obtained models using the C-indices of the development and validation datasets.
Results: Overall, 10 and 21 variables were identified for mortality and morbidity, respectively. C-indices of mortality were 0.940 and 0.924 in the development and validation datasets, respectively. C-indices of morbidity were 0.832 and 0.830 in the development and validation datasets, respectively.
Conclusions: Based on the NCD-P data, we developed satisfactory risk models for mortality and morbidity prediction in major pediatric surgeries.
Level of evidence: Level I (Prognosis Study).

The goal of surgery is to achieve an optimal surgical outcome while minimizing the chances of complications. To achieve this essential goal, every surgeon has to keep his/her training and education up to date, which requires serious commitment [1]; an efficient, systematic effort to reduce complications is also necessary. Beyond individual effort, however, building a risk-adjusted estimation model for mortality and morbidity that systematically guides operations is one of the crucial steps. The aforementioned factors are the essentials needed to maintain and improve the level of our healthcare. [2].

The American College of Surgeons National Surgical Quality Improvement Program-Pediatric (NSQIP-P) is one of the most successful projects to provide an objective assessment of outcomes of pediatric surgeries. Risk models of mortality and morbidity of 382 procedures have already been constructed, providing a web-based risk calculator that is available to the public [3]. In the clinical settings, the risk-adjusted outcomes are useful for preoperative counseling, quality assessment of healthcare services, and medical research as a whole.

* Corresponding author at: Department of Pediatric Surgery, Chiba University Graduate School of Medicine, 1-8-1 Inohana, Chuo-ku, Chiba, 260-8677, Japan. Tel.: +81 43 2227171; fax: +81 43 2262366.
E-mail address: kta@cc.rim.or.jp (K. Terui).

https://doi.org/10.1016/j.jpedsurg.2020.03.031
0022-3468/

In 2011, a nationwide project to register surgical procedures performed in Japan, named the National Clinical Database (NCD), commenced [4,5]. This registry covered more than 95% of general surgeries, and more than 1,700,000 surgical procedures were registered from more than 5000 hospitals in 2016 [6]. NCD registry continues to store patient data annually and provides criteria for the board certification systems for surgeons in many fields. Based on this nationwide registry, the National Clinical Database-Pediatric (NCD-P), which is a more detailed database focusing on pediatric surgeries, commenced in 2015.

The present study, therefore, aimed to construct new risk models for pediatric surgeries using the NCD-P data. Our subsequent goal is to develop and provide a risk calculator system. And, ultimately, we are aiming to improve the quality of our health care by risk-adjusted quality assessment. Risk-adjusted comparative study of hospital volume, regional biases and certification of surgeons are under consideration.

The key feature of the NCD-P was that patient variables and definitions were designed almost identical to those used in the NSQIP-P. Thus, it is possible to validate the risk models of the NSQIP-P using data of the Japanese patients and to also compare these models with our local risk models. In the present study, as the first step towards achieving the goal, 12 major pediatric surgical procedures were selected on the basis of requirements of the board certification system for pediatric surgeons by the Japanese Society of Pediatric Surgeons (JSPS). To our knowledge, this is the first report on the mortality and morbidity for major pediatric surgeries based on data from the NCD-P, the nationwide clinical database in Japan.

1. Method

1.1. Data collection

NCD is a nationwide project to register surgical procedures performed in Japan [5,6]. The service and study design were approved by the Research Ethics Committee of the Graduate School of Medicine, Tokyo University (no. 3111) and also by the Japan Surgical Society Ethics Committee. The requirement to obtain signed informed consent was waived as de-identified data were used as well as opt-out procedures were well-established. Details of the project were published on the official NCD website, and individuals had the right to decline participation in any study originated from NCD. The studies with NCD were performed in accordance with the principles of the Declaration of Helsinki and the Ethical Guidelines for Medical and Health Research Involving Human Subjects.

NCD provides a web-based data management system based on a surgeon's self-reported data. To ensure traceability of data, data is entered by a specific data manager designated by each institution and approved by responsible chairperson. Data managers normally reviewed medical records to complete the entry form and occasionally contact surgeons to make an inquiry concerned surgical procedures including following data after discharge. If the patients do not visit the hospital after discharge, data managers should contact their families or the hospital in which the patients were transferred to access 30-day outcomes. A person in charge of data managing varies in each hospital and is practically played by a medical clerk or a physician. To maintain the quality of the entry data, data managers participate in training programs as well as an E-learning system on the NCD Website. Abundant Frequently Asked Questions about data entry are released on the NCD Website and updated sequentially. Furthermore, Accuracy of data was consecutively validated through an audit process that inspects randomly chosen institutions.

The National Clinical Database has a layered structure. All types of surgical procedures are to be entered via simple data entry; this constitutes the first layer. The second layer is for procedures in more specialized fields (i.e., gastroenterological, cardiovascular, respiratory, etc.) with more detailed data entry. The NCD-P is located on the second layer for pediatric surgeries, which was designed by the JSPS. The NCD-P contains data of high-level pediatric surgeries, neonatal surgeries, and pediatric endoscopic surgeries. In the current study, we focused on the following 12 major pediatric surgical procedures selected on the basis of the board certification system by the JSPS: tracheoplasty, pneumonectomy, fundoplication, and total/subtotal excision of malignant tumor, and surgeries for Hirschsprung disease, anorectal malformation, biliary atresia, choledocal cyst, midgut volvulus, funnel chest, gastrointestinal perforation, and intestinal obstruction. Definitions of the 12 surgical procedures are listed in Table 1. Each procedure contains all approaches, including laparotomy, thoracotomy, and laparoscopic and thoracoscopic surgeries. Pneumonectomy includes lobectomy (one lobe or more lobes) and segmentectomy but does not include wedge resection or bullectomy. Surgery for anorectal malformation does not include surgeries for low-type anorectal malformations. Surgery for gastrointestinal perforation includes percutaneous drainage, surgical drainage, enterostomy omentopexy, serosal patch, enterorrhaphy, open abdominal management, and gastric/intestinal resection but excludes appendicular perforation associated with appendicitis. The sur-

Table 1
Definitions of the 12 major surgical procedures.

Surgery for gastrointestinal perforation	Surgery for perforation of the stomach, the small bowel, and the colon. This included percutaneous drainage, surgical drainage, enterostomy, omentopexy, serosal patch, enterorrhaphy, open abdominal management, and gastric/intestinal resection. Appendicular perforation associated with appendicitis was not included.
Fundoplication	Fundoplication for gastroesophageal reflux disease, hiatus hernia, esophageal stenosis, and esophageal achalasia.
Total/sub-total resection of malignant tumor	Total/sub-total resection of malignant tumor. Biopsies were not included.
Surgery for Hirschsprung disease	Pull-through procedures for Hirschsprung disease including the Soave, Swenson, Duhamel, and Martin procedures.
Surgery for funnel chest	Surgery for funnel chest, including the Nuss procedure, sterno-costal elevation, and sternal turnover.
Surgery for anorectal malformation	Anorectoplasty for high or intermediate type of anorectal malformation. Surgeries for low-type anorectal malformations were not included.
Surgery for intestinal obstruction	Surgery for mechanical intestinal obstruction due to such conditions as adhesion, internal hernia, torsion, intussusception, and tumorous lesion. This included release of intestinal obstruction, adhesiolysis, intestinal resection with or without anastomosis, stricture plasty, enterostomy, intestinal bypass, and stent placement.
Surgery for midgut volvulus	Surgery for midgut volvulus with or without malrotation. This included untwisting the volvulus, the Ladd procedure, resection of necrotic intestine, enterostomy, and drainage.
Surgery for choledocal cyst	Hepaticoenterostomy for choledocal cyst or pancreaticobiliary maljunction without dilation of the bile duct.
Pneumonectomy	This included lobectomy of one or more lobe(s) with or without tracheoplasty, and segmentectomy. Wedge resection and bullectomy were not included.
Surgery for biliary atresia	Kasai portoenterostomy or hepaticoenterostomy for biliary atresia.
Tracheoplasty	Tracheoplasty with or without a graft.

Each surgical procedure contains all the approaches, including laparotomy, thoracotomy, and laparoscopic and thoracoscopic surgery.

gery for intestinal obstruction includes release of intestinal obstruction, adhesiolysis, and intestinal resection with or without anastomosis, stricture plasty, enterostomy, intestinal bypass, and stent placement. These surgical procedures were performed by various kinds of surgeons including certified pediatric surgeons, general surgeons and also trainees.

The variables chosen for the NCD-P were almost identical to those used in the NSQIP-P [7]. Discrepancies between outcomes and definitions of NSQIP-P and NCD-P are listed in Supplementary Table 1. We are discussing closing the gap between NCD-P and NSQIP-P; deep vein thrombosis (DVT) thrombophlebitis is one of the candidates of newly added variables, for instance.

1.2. Patient selection

In the NCD-P records, data from patients who underwent the 12 surgical procedures described above from January 1, 2015 through December 31, 2016, were used. Records in 2015 were used as the development dataset for the risk model construction, and those in 2016 were used as the validation dataset. Patients who refused for their data to be used were excluded from this study. If one patient underwent two or more surgeries simultaneously, the one that was considered major by the performing surgeon was adopted.

1.3. Pre- and perioperative variables

The potential independent variables for the models included patient demographics and preoperative status. Demographic variables, such as body weight and height at surgery and sex, were selected. We also considered whether patients were transferred to hospitals by ambulance. The preoperative status was assessed based on the following: general status (assessed via The American Society of Anesthesiologists Physical Status Classification, congenital anomaly and cardiopulmonary resuscitation), cardiovascular status (assessed via the level of cardiac risk for surgery, congenital heart anomaly, previous cardiac surgery, and dependency on inotrope), respiratory status (assessed via dependence on oxygen, dependence on ventilator within 48 h preoperatively, pneumonia, asthma, chronic lung disease, tracheal malformation, and preexisting tracheostomy), renal status (assessed via renal failure and dialysis), hematological status (assessed via hemostatic disorder and preoperative blood transfusion), neurological status (assessed via developmental delay, cerebral palsy, neuromuscular disorder, congenital central nervous system abnormality, cerebrovascular disorder, and coma), oncological status (assessed via chemotherapy, radiotherapy and bone marrow transplantation), chronic disease (assessed via diabetes, metabolic disease, and hematological disorder), nutritional status (assessed via need for supplemental nutrition and weight loss over 10%), other comorbidities (such as digestive disease, hepatopancreaticobiliary disease, and liver dysfunction), history of organ transplantation, chronic steroid or immune suppressor use, sepsis, history of another surgery within 30 days, and preexisting open wound. We also considered whether the surgery was emergent, unexpected re-operation, simultaneous, oncologic, endoscopic, or involved an anesthesiologist. Surgical procedures were used as risk factors if necessary.

1.4. Endpoints

The primary outcomes of this study were 30-day mortality and composite morbidity. Mortality and morbidity were assessed at postoperative day 30, regardless of whether the patient had been discharged after their initial admission. Composite morbidity was defined as any new-onset incidence of complications within 30 days after surgery as follows: superficial incisional, deep incisional or organ/space surgical site infection (SSI), wound disruption, anastomotic dehiscence, pneumonia, unplanned intubation, pulmonary embolism, progressive renal insufficiency, urinary tract infection, cerebral vascular accident, intraventricular bleeding, coma, peripheral nerve injury, cardiac arrest requiring resuscitation, use of extracorporeal membrane oxygenation, transfusion of >25 ml/kg of red blood cells, systemic sepsis, catheter-related blood stream infection, liver dysfunction, unexpected reoperation, and unexpected re-admission.

1.5. Statistical analysis

We developed two types of risk models (i.e., one for 30-day mortality and the other for 30-day composite morbidity) with the 2015 dataset. To develop each model, we conducted a multivariate logistic regression analysis using all the potential independent variables (as described in the 1.3 “Pre- and perioperative variables” section) in the model and a forward-backward stepwise selection method with Akaike's Information Criterion (AIC) [8]. We chose the model that minimized AIC to predict the occurrence of the primary outcomes. The stepAIC function in the MASS package of R [9] was used for this procedure. We calculated the area under the curve (AUC), i.e., *C-index*, for each of the risk models to assess the models' discriminatory power.

Subsequently, we evaluated the generalization ability of the best risk models with the 2016 dataset. We calculated the *C-index* for all the best models to predict 30-day mortality and for the 30-day composite morbidity with this dataset. The Hosmer-Lemeshow test was conducted to assess the calibration of the risk models. We calculated the Hosmer-Lemeshow statistic of each risk model using 2015 as well as 2016 datasets. In addition, the calibration curve of each risk model was plotted. The hoslem.test function in the ResourceSelection package was used to conduct the Hosmer-Lemeshow test and the calibrate.plot function in the gbm package was used to plot the calibration curves.

All p values were two tailed and those under 0.05 were considered statistically significant. All statistical analyses were conducted using R.

2. Results

2.1. Demographic and clinical characteristics

A total of 6522 patients who had undergone 12 selected surgical procedures of the NCD-P in 2015 and 2016 were included in this study. The development dataset included 3057 records from 227 institutions, and the validation dataset included 3465 records from 244 institutions. In each cohort, 20 freestanding children's hospitals were included. The patients' demographic and clinical characteristics are shown in Table 2. The percentage of neonates (aged ≤28 days) at surgery were 19% and 17% in the development and validation datasets, respectively.

2.2. Outcome rates

Table 3 shows the mortality and morbidity of 12 surgical procedures. Overall 30-day mortality rate and composite morbidity rate in the development dataset were 1.5% and 21.3%, respectively. The highest mortality was seen in a surgery for gastrointestinal perforation (5.2% in 2015, 5.6% in 2016). The highest morbidity was also seen in a surgery for gastrointestinal perforation (43.2% in 2015, 43.2% in 2016).

Table 2
Demographic data.

		2015		2016	
		n	(%)	n	(%)
No. of institutions		227	-	244	-
All cases		3057	(100)	3465	(100)
Age at surgery	≤28 days	587	(19)	597	(17)
	29 days – <1 year	909	(30)	893	(26)
	1–<6 years	800	(26)	833	(24)
	6–<16 years	761	(25)	1139	(33)
Sex	Male	1725	(56)	2143	(62)
	Female	1332	(44)	1322	(38)

Table 3
Mortality and morbidity of 12 surgical procedures.

Surgical procedure	2015				2016			
	n		Mortality n (%)	Morbidity n (%)	n		Mortality n (%)	Morbidity n (%)
	Total	Per hospital			Total	Per hospital		
Surgery for gastrointestinal perforation	676	0 (0–16)	35 (5.2%)	292 (43.2%)	709	2 (0–17)	40 (5.6%)	306 (43.2%)
Fundoplication	451	0 (0–27)	1 (0.2%)	66 (14.6%)	599	1 (0–23)	2 (0.3%)	88 (14.7%)
Total/sub-total resection of malignant tumor	315	0 (0–28)	1 (0.3%)	57 (18.1%)	430	0 (0–60)	1 (0.2%)	68 (15.8%)
Surgery for Hirschsprung disease	223	0 (0–9)	0 (0.0%)	20 (9.0%)	233	0 (0–9)	0 (0.0%)	31 (13.3%)
Surgery for funnel chest	220	0 (0–36)	0 (0.0%)	7 (3.2%)	268	0 (0–80)	0 (0.0%)	8 (3.0%)
Surgery for anorectal malformation	220	2 (0–14)	0 (0.0%)	40 (18.2%)	216	0 (0–11)	0 (0.0%)	37 (17.1%)
Surgery for intestinal obstruction	207	0 (0–12)	3 (1.4%)	42 (20.3%)	203	0 (0–9)	3 (1.5%)	40 (19.7%)
Surgery for midgut volvulus	191	0 (0–7)	1 (0.5%)	15 (7.9%)	230	0 (0–8)	1 (0.4%)	40 (17.4%)
Surgery for choledocal cyst	187	0 (0–8)	0 (0.0%)	21 (11.2%)	194	0 (0–8)	0 (0.0%)	19 (9.8%)
Pneumonectomy	180	0 (0–16)	3 (1.7%)	23 (12.8%)	214	0 (0–11)	1 (0.5%)	23 (10.7%)
Surgery for biliary atresia	147	0 (0–9)	0 (0.0%)	52 (35.4%)	116	0 (0–7)	0 (0.0%)	47 (40.5%)
Tracheoplasty	40	0 (0–10)	1 (2.5%)	16 (40.0%)	53	0 (0–12)	1 (1.9%)	18 (34.0%)
Total	3057		45 (1.5%)	651 (21.3%)	3465		49 (1.4%)	725 (20.9%)

Top 10 morbidities during 2015 and 2016 were transfusion of >25 ml/kg of red blood cells (7.8%), liver dysfunction (4.3%), superficial incisional SSI (3.8%), unexpected reoperation (3.5%), systemic sepsis (2.9%), wound disruption (2.7%), pneumonia (2.0%), organ/space SSI (1.7%), unplanned intubation (1.6%), catheter-related blood stream infection (1.4%) (Supplementary Table 2).

2.3. Model results

Selected variables for 30-day mortality and composite morbidity are given in Tables 4 and 5, respectively. Odds ratio for 30-day composite morbidity of each surgical procedure are showed in Supplementary Table 3.

The following 10 variables were selected for the best risk model to predict 30-day mortality: cardiac risk factor, dependence of respirator within 48 h, immune suppressor use, dependence of hemodialysis, emergent surgery, renal failure, coma over 24 h, body weight at surgery <1.5 kg, need for supplemental nutrition, and unexpected reoperative surgery. The *C-index* for the mortality model was 0.940 (Table 6).

The following 21 variables were selected for the best risk model to predict 30-day composite morbidity: dependence of respirator within 48 h, body weight at surgery <1.5 kg, ASA, sepsis, chronic lung disease, chemotherapy within 30 days, bone marrow transfer, hepatopancreaticobiliary disease, dependence of inotrope, involvement of anesthesiologist, dependence of oxygen, history of blood transfusion within 48 h, transportation by ambulance, gastrointestinal disease, sex, emergent surgery, history of another surgery within 30 days, congenital anomaly, liver dysfunction, dependence of hemodialysis, and preoperative tracheostomy. To improve the ability of the model, surgical procedures were also used as risk factors. Odds ratio of each procedure for the surgery for gastrointestinal perforation was showed in Supplementary Table 3. The *C-index* for the composite morbidity model was 0.832 (Table 6).

2.4. Model performance

To assess discriminatory power of the risk models for the validation dataset, we calculated the *C-index* of each of the best models to predict

Table 4
Selected variables for 30-day mortality model.

Variables		Mortality n (%)	Morbidity n (%)	Odds ratio (95% confidence interval)	
Cardiac risk factor	High	8/41 (20)	24/41 (59)	13.8	(4.3–41.0)
	Intermediate	5/91 (5)	37/91 (41)	1.6	(0.4–4.5)
	Low	12/221 (5)	90/221 (41)	1.6	(0.7–3.6)
	None	20/2704 (1)	500/2704 (18)	-	
Dependence of respirator within 48 h		39/444 (9)	225/444 (51)	8.9	(3.3–26.9)
Immune suppressor use		3/29 (10)	12/29 (41)	7.1	(10. – 34.9)
Dependence of hemodialysis		6/15 (40)	12/15 (80)	4.9	(1.0–23.8)
Emergent surgery		38/941 (4)	318/941 (34)	3.9	(1.5–11.2)
Renal failure		11/61 (18)	27/61 (44)	2.9	(0.9–8.3)
Coma over 24 h		7/28 (25)	14/28 (50)	2.9	(0.8–9.5)
Need for supplemental nutrition		23/555 (4)	179/555 (32)	2.0	(1.0–3.9)
Body weight at surgery <1.5 kg		24/241 (10)	146/241 (61)	0.5	(0.2–1.1)
Unexpected reoperative surgery		9/284 (3)	100/284 (35)	0.5	(0.2–1.1)

Table 5
Selected variables for 30-day morbidity model.

Variables		Mortality n (%)	Morbidity n (%)	Odds ratio (95% confidence interval)	
Dependence of hemodialysis		6/15 (40)	12/15 (80)	3.4	(0.9–17)
ASA	5	16/76 (21)	58/76 (76)	3.4	(1.7–6.7)
	4	8/118 (7)	63/118 (53)	1.8	(1.1–2.9)
	3	13/555 (2)	179/555 (32)	1.5	(1.1–2.0)
	2	6/1100 (1)	202/1100 (18)	1.1	(0.8–1.4)
	1	2/1208 (0)	149/1208 (12)	-	
Sepsis	Septic shock	10/47 (21)	39/47 (83)	3.0	(1.3–7.5)
	Sepsis	3/53 (6)	25/53 (47)	0.7	(0.3–1.3)
	SIRS	5/68 (7)	40/68 (59)	1.8	(1.0–3.3)
	None	27/2889 (1)	547/2889 (19)	-	
Chronic lung disease		7/118 (6)	61/118 (52)	2.1	(1.3–3.3)
Chemotherapy within 30 days		0/116 (0)	29/116 (25)	1.7	(1.0–3.0)
Bone marrow transfer		0/39 (0)	10/39 (26)	1.7	(0.7–3.6)
Hepatopancreaticobiliary disease		0/319 (0)	87/319 (27)	1.6	(1.0–2.5)
Dependence of inotrope		23/167 (14)	111/167 (66)	1.6	(1.0–2.5)
Involvement of anesthesiologist		36/2948 (1)	612/2948 (21)	1.5	(0.9–2.6)
Dependence of oxygen		38/566 (7)	258/566 (46)	1.4	(1.0–2.0)
History of blood transfusion within 48 h		13/159 (8)	89/159 (56)	1.4	(0.9–2.1)
Gastrointestinal disease		20/843 (2)	233/843 (28)	1.3	(1.0–1.7)
Ambulance transport		12/510 (2)	161/510 (32)	1.3	(1.0–1.7)
Female sex		20/1332 (2)	315/1332 (24)	1.2	(1.0–1.5)
Dependence of respirator within 48 h		39/444 (9)	225/444 (51)	1.2	(0.8–1.8)
History of another surgery within 30 days		12/279 (4)	107/279 (38)	1.2	(0.9–1.6)
Congenital anomaly		6/459 (1)	108/459 (24)	1.2	(0.9–1.6)
Emergent surgery		38/941 (4)	318/941 (34)	1.1	(0.8–1.6)
Liver dysfunction		5/301 (2)	101/301 (34)	1.1	(0.7–1.6)
Preoperative tracheostomy		2/161 (1)	31/161 (19)	0.8	(0.5–1.2)
Body weight at surgery <1.5 kg		24/241 (10)	146/241 (61)	0.6	(0.4–0.9)

Abbreviations: ASA, American Society of Anesthesiologists class; SIRS, systemic inflammatory response syndrome.

30-day mortality and for 30-day composite morbidity using 2016 dataset. The *C-index* for mortality was 0.924, and that for composite morbidity was 0.830 (Table 6).

The Hosmer-Lemeshow test showed that statistical significance was only seen in the 30-day morbidity model (Table 6). However, calibration plots show that there was a tendency for the both of 30-day mortality and 30-day composite morbidity models to overestimate risk (Fig. 1).

3. Discussion

NSQIP-Pediatric is one of the most successful projects to provide a risk-adjusted quality assessment of pediatric surgical procedures [3]. Data of NSQIP-Pediatric has been collected systematically in high-volume centers in the US, resulting in very reliable database. However, medical systems and patients' backgrounds are different in Japan from those provided in the NSQIP-P project. Centralization has not been established well, and real high-volume centers are quite few in Japan. It remains unclear, therefore, as to whether or not NSQIP-P can be applied to our medical systems different from the US. Actually, a comparative study of adult gastroenterological surgery between the US and Japan showed many differences including patient background, comorbidities, and outcomes. The researchers also compared risk models in both sides and concluded that the local risk models could accurately predict the outcome [10]. The purpose of the present study comes from the above question. It is to newly construct a risk model for pediatric surgical procedures by using Japanese nation-wide database, NCD-P.

Table 6
Model Statistics for the 30-day mortality and morbidity model

	Development dataset		Validation dataset	
	C-index	Hosmer-Lemeshow (p Value)	C-index	Hosmer-Lemeshow (p Value)
30-day mortality	0.940	0.248	0.924	0.280
30-day morbidity	0.832	0.407	0.830	< 0.001

The NCD-P reflects the reality of pediatric surgery in Japan most accurately because it is the biggest and most thorough database available in Japan, comprising data from most of the institutions that perform pediatric surgeries across the country. In the development dataset, 3057 children who underwent surgery in 227 institutes throughout Japan were analyzed. The 30-day mortality and composite morbidity in the development dataset were 1.5% and 21.3%, respectively; these rates are higher than those for the NSQIP-P model, which were 0.3% and 5.9%, respectively [3]. These differences in outcome probably resulted from the difference in procedures in the studies; the NSQIP-P involved 382 procedures, including general surgery, cardiothoracic, neurosurgery, orthopedics, otolaryngology, gynecology, urology, and plastic surgery [3]. The NCD-P counterpart targeted the 12 surgical procedures that a surgeon must be versed in to be a JSPS-certified pediatric surgeon.

The 30-day mortality and morbidity models included 10 and 21 variables, respectively. The most powerful preoperative risk factors for mortality and morbidity were "preoperative cardiac risk: high" and "dependence of ventilator use within 48 h", respectively. The selected risk factors were nearly identical to known risk factors for postoperative mortality [12,13].

The C-indices for mortality and morbidity in the validation dataset were 0.924 and 0.830, respectively. The C-indices for mortality using the NSQIP-P model were 0.978 and 0.806, respectively, for the development and validation datasets, and the average C-indices for morbidity using the NSQIP-P model were 0.912 and 0.869, respectively, for the development and validation datasets [3]. The C-indices for the mortality of the Preoperative Complication Score and the Overall Complication Score were 0.740 and 0.767, respectively [14]. Therefore, among the known predictive models, the NSQIP-P had the best predictive ability for the prognosis of pediatric surgeries; the predictive ability of the NCD-P was identical to that of the NSQIP-P.

The key point of the present study is that the variables were selected from those used in the NSQIP-P. We can potentially compare not only the risk models between the two registry systems, but also the medical

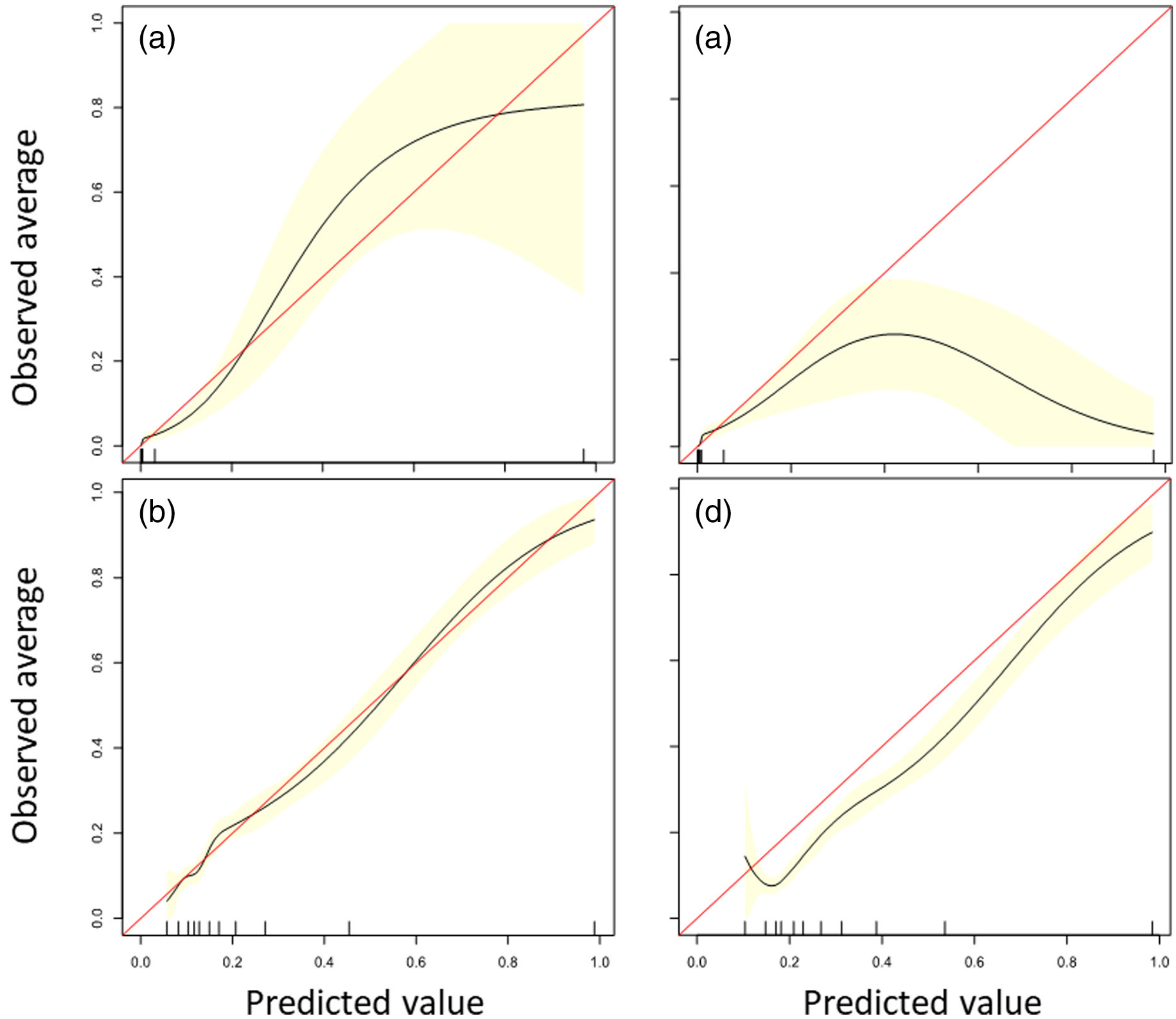

Fig. 1. Calibration curve for 30-day mortality and morbidity model. (a) mortality in development dataset, (b) mortality in validation dataset, (c) morbidity in development dataset, (d) morbidity in validation dataset.

systems at work behind the data between Japan and the US. This may be useful to maintain and improve the level of healthcare systems in both countries.

This study has several limitations that need to be addressed in the future studies. First, data entry of the NCD-P varies in each hospital. Although training programs for the data managers have been set up, incorrect data entry may have occurred. A systematic audit is underway to assure and maintain the quality of data. Second, we targeted only 12 surgical procedures in the study. To address and overcome this limitation, we have added another set of 26 procedures to the NCD-P since 2017, and the data will be analyzed in due course. Third, the reported mortality and morbidity rates in our study could have been influenced by factors not included in the list of our variables, such as hospital volume, institutional and individual experience, regional biases, and patients' conditions. These factors should be assessed by risk-adjusted comparisons in the future. Fourth, the study population was almost limited to a single racial composition. Therefore, our results should be validated by a cohort of other ethnicities for broader usage. Fifth, the mortality model had insufficiently fit in the calibration plot with the validation dataset. However, with such extremely low rate of mortality, this overestimation has little effect on actual use of the model. The morbidity model was significant in Hosmer-Lemeshow test in validation dataset. According to the calibration plot, the morbidity model overestimates the risk by approximately 10%. Further adjustment of the model is needed. Sixth, the 12 different surgical procedures were combined to construct a single model in the present analysis. This combination of surgical procedures enabled us to develop better risk models in 2 difficult situations: One was small number of each surgical procedure even if national data was used. The other was low incidence of adverse event. These 2 disadvantages made it difficult to create a good risk model. Consequently, in the model of 30-day mortality, the prognostic ability was satisfactory not by using information of surgical procedures. This means that whatever surgery a patient undergoes, the risk of mortality can be predicted only by using preoperative variables. On the other hand, in the model of 30-day composite morbidity, surgical procedures needed to be used as risk factors for better prognostic ability. Each complication was also grouped together into the composite morbidity. These analytical procedures made complicated situations simpler to produce more robust models, but the specificity of each surgery as well as each complication may have been lost. Therefore, we are preparing more informative models that account specific complications for specific procedures. Seventh, we did not assess procedure-associated risk. In Japan, a case-mix system called DPC(Diagnosis Procedure Combination) was established in 2012 and has been used as a basement of the payment system [15]. However, NCD has not been directly connected with DPC until now. The collaboration between NCD and DPC is one of our future tasks. Eighth, the surgical procedures were performed by various kinds of surgeons including certified pediatric surgeons, general surgeons, and trainees. The participation of trainees has been reported as a risk factor [16]. However, the type of surgeon was not included in the risk factors in this study mainly because we could not access the surgeon type of each doctor due to personal information protection. Assessment of the risk of the surgeon type is definitely our future task.

4. Conclusion

We report the first risk stratification study of pediatric surgery based on the Japanese nationwide database, the NCD-P. By analyzing 3057 patients throughout Japan, we found that the 30-day mortality and composite morbidity rates were 1.5% and 21.3%, respectively. We also developed satisfactory risk models for mortality and morbidity, which should be useful in preoperative counseling and also contribute to an improved quality control in surgical practice. A web-based risk calculator is now under construction.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpedsurg.2020.03.031.

Acknowledgments

We thank all the data managers and hospitals participating in the NCD-P project for their continued efforts in entering the data.

Funding

This work was partly supported by JSPS KAKENHI Grant Number 17H04146.

References

[1] Caty MG, Cowles RA. Complications in pediatric surgery. Preface Semin Pediatr Surg. 2016;25:337. https://doi.org/10.1053/j.sempedsurg.2016.10.001.

[2] Porter ME. What is value in health care? N Engl J Med. 2010;363:2477–81. https://doi.org/10.1056/NEJMp1011024.

[3] Kraemer K, Cohen ME, Liu Y, et al. Development and evaluation of the American College of Surgeons NSQIP pediatric surgical risk calculator. J Am Coll Surg. 2016;223:685–93. https://doi.org/10.1016/j.jamcollsurg.2016.08.542.

[4] Miyata H, Gotoh M, Hashimoto H, et al. Challenges and prospects of a clinical database linked to the board certification system. Surg Today. 2014;44:1991–9. https://doi.org/10.1007/s00595-013-0802-3.

[5] Anazawa T, Miyata H, Gotoh M. Cancer registries in Japan: National Clinical Database and site-specific cancer registries. Int J Clin Oncol. 2015;20:5–10. https://doi.org/10.1007/s10147-014-0757-4.

[6] National Clinical Database. http://www.ncd.or.jp/; 2019. [Accessed 24 May 2019].

[7] Shiloach M, Frencher Jr SK, Steeger JE, et al. Toward robust information: data quality and inter-rater reliability in the American College of Surgeons National Surgical Quality Improvement Program. J Am Coll Surg. 2010;210:6–16. https://doi.org/10.1016/j.jamcollsurg.2009.09.031.

[8] Akaike H. Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K, Kitagawa G, editors. Selected papers of Hirotugu Akaike. New York: Springer; 1998. p. 199–213.

[9] R Core Team. R: A language and environment for statistical computing. https://www.R-project.org/; 2018. [Accessed 24 May 2019].

[10] Anazawa T, Paruch JL, Miyata H, et al. Comparison of National Operative Mortality in gastroenterological surgery using web-based prospective data entry systems. Medicine. 2015;94:e2194. https://doi.org/10.1097/MD.0000000000002194.

[12] Weinberg AC, Huang L, Jiang H, et al. Perioperative risk factors for major complications in pediatric surgery: a study in surgical risk assessment for children. J Am Coll Surg. 2011;212:768–78. https://doi.org/10.1016/j.jamcollsurg.2011.02.006.

[13] Lian C, Xie Z, Wang Z, et al. Pediatric preoperative risk factors to predict postoperative ICU admission and death from a multicenter retrospective study. Paediatr Anaesth. 2016;26:637–43. https://doi.org/10.1111/pan.12905.

[14] Wood G, Barayan G, Sanchez DC, et al. Validation of the pediatric surgical risk assessment scoring system. J Pediatr Surg. 2013;48:2017–21. https://doi.org/10.1016/j.jpedsurg.2013.04.017.

[15] Matsuda S. Development and use of the Japanese case-mix system. Eurohealth. 2008;14:25–30.

[16] Lederhuber H, Hanßke B, Dahlstrand U. Impact of trainee participation on inguinal hernia repair outcome: a study based on the Swedish hernia register. Ann Surg. 2019. https://doi.org/10.1097/SLA.0000000000003497 [Epub ahead of print].