Esophageal atresia: Prognostic classification revisited

Tatsuya Okamoto, MD,[a] Shigeru Takamizawa, MD,[b] Hiroshi Arai, MD,[a] Yuko Bitoh, MD,[a] Makoto Nakao, MD,[a] Akiko Yokoi, MD,[a] *and* **Eiji Nishijima, MD,[a]** *Kobe and Nagano, Japan*

Background. *Although the Spitz classification is the most widely used prognostic classification for esophageal atresia and/or tracheoesophageal fistula (EA), its discrimination ability remains unclear. We sought to develop a more accurate prognostic classification for EA.*

Methods. *The records of 121 consecutive infants with EA (1980–2005) were reviewed. The independent variables included 6 clinical characteristics, and the dependent variables were survival and mortality. Stepwise logistic regression analysis was used to construct models predicting mortality and create a revised prognostic classification. The discrimination abilities of the revised classification and the Spitz classification were compared using receiver-operating characteristic (ROC) curves.*

Results. *Birth weight and the presence of major cardiac anomalies were significant prognostic factors for mortality, and major cardiac anomalies affected mortality more than birth weight. The ROC curve for birth weight suggested that 2,000 g was an appropriate cutoff point. The Spitz classification was revised as follows: the revised class I (low-risk group) consisted of patients without major cardiac anomalies and birth weight >2,000 g; class II (moderate-risk group) consisted of patients without major cardiac abnormalities and birth weight <2,000 g; class III (relatively high-risk group) consisted of patients with major cardiac anomalies and birth weight >2,000 g; and class IV (high-risk group) consisted of patients with major cardiac anomalies and birth weight <2,000 g. The ROC comparisons showed that the revised classification provided a significant improvement (*P = .049*).*

Conclusion. *This revised classification can improve the stratification of EA patients and be a useful predictor of survival. (Surgery 2009;145:675-81.)*

From the Department of Pediatric Surgery, Kobe Children's Hospital,[a] Kobe; and the Department of Pediatric Surgery, Nagano Children's Hospital,[b] Nagano, Japan

Since Waterston et al[1] introduced a prognostic classification for esophageal atresia and/or tracheoesophageal fistula (EA) in 1962, several prognostic classifications have been proposed. Currently, the Spitz classification, which is based on birth weight and the presence or absence of major cardiac anomalies and was proposed in a review of 357 cases of EA treated between 1980 and 1992,[2] has been recognized as the most commonly used prognostic classification for EA.[3] However, as shown in the original data from Spitz et al[2,4] and the outcomes of series of EA patients based on the Spitz classification,[5-10] the numbers of infants who fulfilled the criteria of Spitz group III (high-risk group) were too small to allow a valid comparison. Furthermore, there have been major advances in the fields of neonatology, pediatric cardiology, and surgery, and the survival of Spitz group I (low-risk group) seems to have been unchanged over the last 2 decades. We hypothesized that the imbalance in the distribution of the number of patients in the high-risk group and the unchanging survival rate of the low-risk group may be the result of some limitations of the classification itself. Therefore, we reviewed the outcome of infants born with EA in our hospital to develop an accurate prognostic classification; the prognostic reliabilities of the 2 classification systems were compared using receiver-operating characteristic (ROC) curve analysis.

PATIENTS AND METHODS

The study involved 121 consecutive EA patients treated at Kobe Children's Hospital between January 1980 and December 2005. This study was done with the approval of the institutional review board (registered no. R20-3). The data collected included the anatomic type of EA, operative interventions, gestational week, weight at birth, presence of major cardiac anomalies, presence of

Accepted for publication January 16, 2009.

Reprint requests: Tatsuya Okamoto, MD, Department of Surgery, Kyoto University Graduate School of Medicine, 54-Kawahara-cho, Sho-goin, Sakyo-ku, Kyoto 606-8507, Japan. E-mail: okamoto_kch@hp.pref.hyogo.jp; tokamoto@mbn.nifty.com.

0039-6060/$ - see front matter

doi:10.1016/j.surg.2009.01.017

Table I. Associated cardiac anomalies (n = 37)

Anomaly	n
Vascular ring	1
PDA	5
ASD	1
VSD	2
PDA + ASD	1
PDA + VSD	7
ASD + VSD	1
PDA + ASD + VSD	1
PA + ASD + VSD	1
Coarctation of the aorta	1
Aortic coarctation complex	2
TA + ACC	1
Total anomalous pulmonary venous return	1
Double outlet right ventricle	3
Tetralogy of fallot	6
Hypoplastic left heart syndrome	1
SV	2

ASD, Atrial septal defect; *PA*, pulmonary atresia; *PDA*, patent ductus arteriosus; *TA*, tricuspid atresia; *VSD*, ventricular septal defect; *SV*, single ventricle.

Table II. Total number of associated anomalies (n = 40) and chromosomal abnormalities (n = 11)

Anorectal anomalies	14
VATER associated	8
Cleft left	5
Duodenal atresia	3
Hydronephrosis	2
Goldenhar syndrome	2
Micrognathia	2
Lagophthalmos	2
Polydactyly	2
Congenital bronchial agenesis	2
Hydrocephalus	1
Congential diaphragmatic hernia	1
Bronchopulmonary foregut malformation	1
Radial defect	1
Genitourinary abnormalities	1
Cerebellar hypoplasia	1
Chromsomal abnormalities	
Trisomy 18	6
Trisomy 21	5

associated anomalies other than cardiac anomalies, presence of chromosomal abnormalities, and patient outcomes. Major cardiac anomalies were defined as either a cyanotic congenital heart anomaly that required palliative or corrective operation, including operative correction for patent ductus arteriosus, or a noncyanotic congenital heart that required medical or operative treatment for cardiac failure.

The outcome of interest was the survival rate of infants with EA; any child discharged from hospital was considered a survivor. Three children who were withdrawn from treatment were included in the survival data calculation.

The Student t test was used to compare continuous variables, and the Fisher exact test or χ^2 test was used to compare categorical variables. To develop a prognostic classification model, variables with a P value <.2 in the univariate analysis were included in the multivariate logistic regression model using a Wald statistic with backward, stepwise selection. The discrimination of the new prognostic classification was quantified and compared with that of the Spitz classification using the area under the ROC curve (AUC). Statistical analysis was performed using SPSS software version 11.0 (SPSS Inc., Chicago, IL), and the AUC comparison was carried using the method described by Hanley and McNeil.[11] $P < .05$ was considered significant.

RESULTS

A total of 121 infants with EA were admitted to our institute over the 26-year period. Anatomic variations of total cases were as follows: EA + distal tracheoesophageal atresia (TEF), 101 (84%); EA + proximal TEF, 1 (1%); EA + distal and proximal TEF, 1 (1%); pure EA, 15 (12%); and H-type EA, 3 (3%). Among the 121 infants, there were 67 males and 54 were females. Average of birth weight of all cases was 2,330 ± 660 g, and the average gestational week was 37 ± 3 weeks. Spitz classification were as follows: 75 were Spitz class I (birth weight >1,500 g, without major cardiac disease; 72 were alive and the survival rate was 96%); 27 were Spitz class II (birth weight <1,500 g, or major cardiac disease; 27 cases were alive and the survival rate was 66%); 5 were Spitz class III (birth weight <1,500 g, and major cardiac disease; 2 cases were alive and the survival rate was 40%). Treatment procedures were primary anastomosis (n = 93), staged repair, (n = 22), and cervical approach of TEF (H-type; n = 3); there were 3 patients in whom treatment was withdrawn. The associated major cardiac anomalies (n = 37), the total number of associated anomalies (n = 40), and the chromosomal abnormalities (n = 11) are listed in Tables I and II.

Table III shows the comparison of the 6 variables: gestational week, birth weight, presence of major cardiac anomalies, presence of trisomy 18, presence of trisomy 21, and presence of associated anomalies. Gestational week (P = .019), birth weight (P < .001), presence of major cardiac anomalies (P < .001), and presence of trisomy 18 (P < .001) were significantly associated with mortality.

After univariate analysis, the variables with a P < .2 were selected for inclusion in the multivariate,

Table III. Comparisons of clinical characteristics between living (n = 101) and dead (n = 20) patients

Variable	*Alive*	*Dead*	P *value*
Gestational age (w)	38 ± 3	36 ± 3	.019
Birth weight (g)	2,420 ± 640	1,860 ± 580	<.001
Cardiac anomalies	22/101 (20%)	15/20 (75%)	<.001
Trisomy 18	1/101	5/20	<.001
Trisomy 21	3/101	2/20	.149
Associated anomalies	32/101	8/20	.470

logistic analysis. Presence of major cardiac anomalies was associated with mortality (P < .001; odds ratio [OR],12.414; 95% confidence interval [CI], 3.4–45.6). Lower birth weight was associated with increased mortality (P = .016; OR, 0.987; 95% CI, 0.97–0.99), but this trend was not as strong as the presence of major cardiac anomalies. The multivariate analysis could not show the significant relation between the presence of trisomy 18 and mortality (P = .063; OR, 11.536; 95% CI, 0.87–152.68).

Revised classification. Based on these results, the Spitz classification was revised. Because the presence of major cardiac anomalies was a stronger predictor of survival than low birth weight, the patients were first classified into 2 groups according to the presence or absence of major cardiac anomalies. Then, the 2 groups were further classified by birth weight. The ROC curve for birth weight and the survival rate (Fig 1) was used to identify the cutoff point. According to the ROC curve for birth weight, the optimal cutoff value was about 2,100 g (the intersection point between the ROC curve and the diagonal line drawn from the left upper corner to the right lower corner). To simplify the classification, however, 2,000 g was chosen as the cutoff value for birth weight.

The result was the following revised classification: class I (low-risk group), which consisted of patients without major cardiac anomalies and a birth weight >2,000 g; class II (moderate-risk group), which consisted of patients without major cardiac abnormalities and birth weight <2,000 g; class III (relatively high-risk), which consisted of patients with major cardiac anomalies and birth weight >2,000 g; and class IV (high-risk group), which consisted of patients with major cardiac anomalies and birth weight <2,000 g. The revised classes and the corresponding survival rates are shown in Figure 2. The survival rates were 100% for class I (n = 58), 82% for class II (n = 27), 72% for class III (n = 25), and 27% for class IV (n = 11), which allowed good differentiation between the classes (χ^2 = 39.1; P < .001).

The discrimination abilities of this revised classification and the Spitz classification for mortality

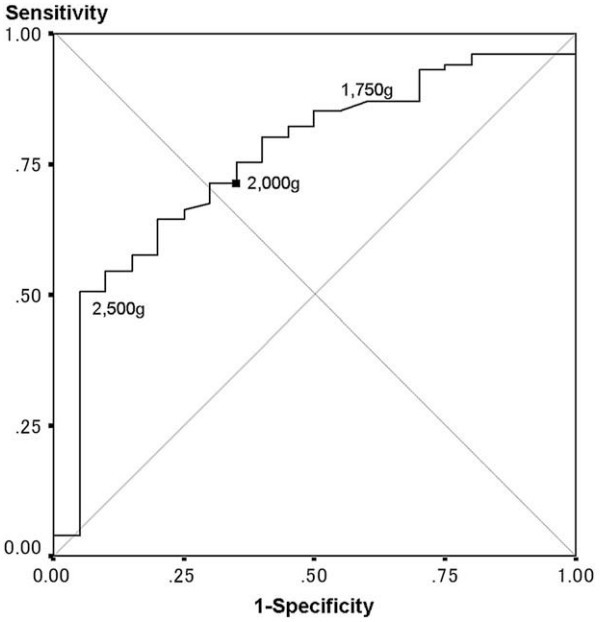

Fig 1. ROC curves for birth weight. The ROC curve shows an optimal cutoff point at about 2,100 g (the intersection point between the ROC curve and the diagonal line drawn from the left upper corner to the right lower corner). The value of 2,000 g was chosen as the cutoff point for birth weight to simplify the classification

were quantified using ROC curves (Fig 3). The AUC of the revised classification was better than that of the Spitz classification (0.873 [95% CI, 0.81–0.94] vs 0.795 [95% CI, 0.69–0.90], respectively; P = .049).

Furthermore, to examine the difference in discrimination of the 2 classifications by time period, the patients were divided in 2 groups according to two 13-year time periods (1980–1992 and 1993–2005). Table IV compares the survival rate for the classes of each classification during the 2 time periods. There was a slight, although not significantly, different improvement in survival for Spitz classes II and III patients, and for the revised classes II, III, and IV patients. The discriminations of both classifications for the 2 time periods are shown by ROC curves in Figure 4, *A* and *B*. In the first 13-year time period, both classifications

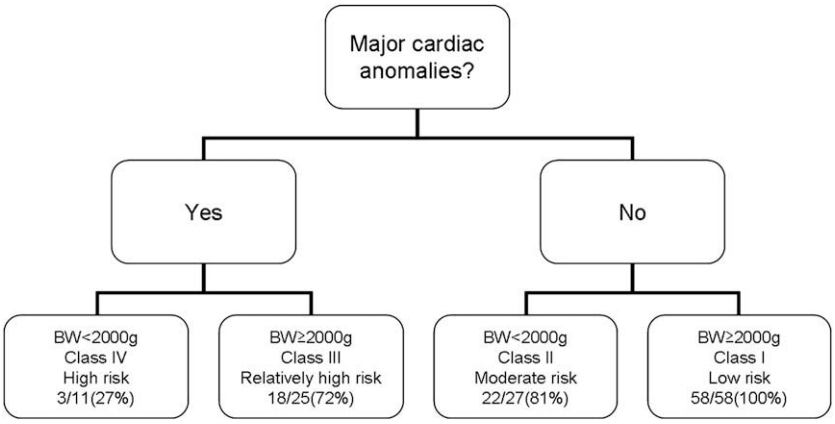

Fig 2. Revised classification and survival rates.

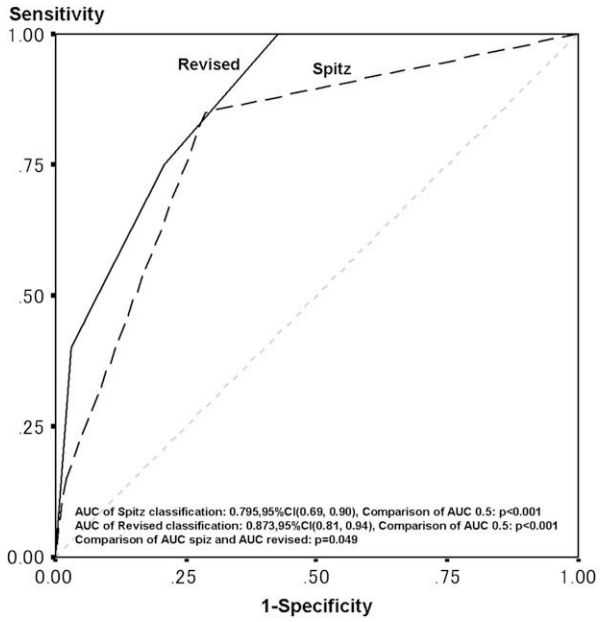

Fig 3. ROC curves for the Spitz classification and the revised classification. The AUC of the revised classification (0.873; 95% CI, 0.81–0.94) is significantly better than that of the Spitz classification (0.795; 95% CI, 0.69–0.90; $P = .049$).

had good discrimination for mortality, whereas the Spitz classification was not a good prognostic predictor during the second 13-year time period.

Finally, the relationships among the classifications, cardiac anomalies, associated anomalies, and causes of death for 20 infants with EA are shown in Table V.

Table IV. Comparison of survival rate of 2 different time periods

	1980–1992 (n = *69*)	*1993–2005* (n = *52*)	P *value*
Spitz class			
I	38/39 (97%)	34/36 (84%)	.60
II	17/28 (60%)	10/13 (76%)	.48
III	0/2 (0%)	2/3 (66%)	.40
Revised class			
I	33/33 (100%)	25/25 (100%)	—
II	8/11 (74%)	14/16 (88%)	.27
III	13/19 (73%)	5/6 (83%)	.36
IV	1/6 (17%)	2/5 (40%)	.54

DISCUSSION

In 1962, Waterston et al[1] proposed a prognostic classification of EA, which was based on birth weight, pneumonia, and associated congenital anomalies. One of the important points of its classification was respiratory state of EA neonate. In response to advances in neonatal care, Poenaru et al[12] suggested that only severe pulmonary dysfunction requiring mechanical ventilation remained a prognostic risk factor in their series, and proposed a new classification based on preoperative ventilator dependence and associated major anomalies. Teich et al[13] reported its superiority to Waterston's classification to identify children at high risk in their series during 1972 to 1996.

With modern neonatal respiratory care, more low birth weight or very low birth weight infants have survived at a greater rate, thus questioning the prognostic significance of the respiratory status of EA as suggested by Waterston's or Poenaru's classifications. Spitz et al[2] proposed their classification based on cardiac anomalies and birth weight in 1994, and currently the Spitz classification is the system used most commonly.

The outcomes of several large series of EA patients have been reported based on the Spitz

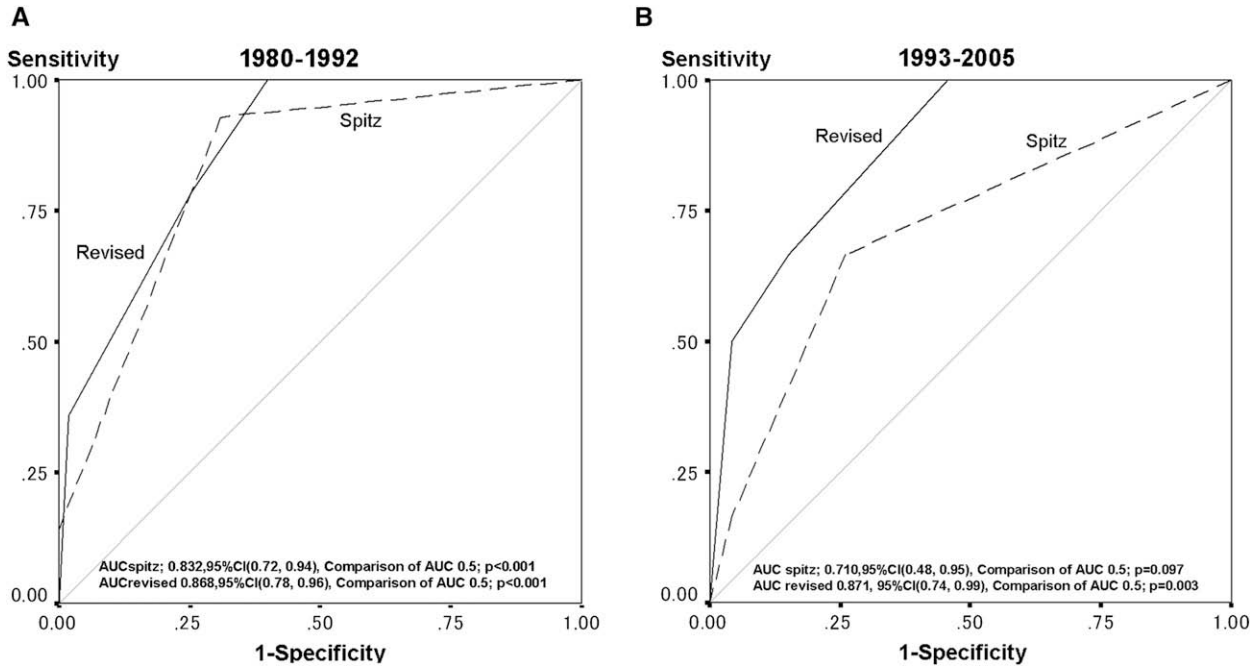

Fig 4. The ROC curves of the Spitz classification and the revised classification for 2 different time periods. (*A*) 1980–1992. (*B*) 1993–2005.

classification,[2,4-10] but the discriminatory ability of this classification has not been measured. The ability of these systems to predict outcome using ROC curves are summarized in Table VI. Although the AUC of the Spitz classification was different from an AUC of 0.5 in all series ($P < .001$), the values varied between 0.861 and 0.657, and the AUC for the cases of all of the cited series was 0.747, which does not seem to be satisfactory for a well-established classification.

As in the study by Spitz et al,[2] the presence of major cardiac anomalies and low birth weight were prognostic predictors for infants with EA in the present study. Two further important findings are suggested by our study: The presence of major cardiac anomalies was a stronger predictor of survival than low birth weight, and the ROC analysis of birth weight suggested that the cutoff value of 1,500 g, which was adopted in the Spitz classification, was not as good as the objective cutoff value of 2,100 g in our study for survival of infants with EA. These results led us to formulate a revised classification.

The first finding seems understandable, because the general survival rate of very low birth weight infants weighing 1,001–1,500 g was already >90% in 1994,[14] whereas the operative mortality risk for cardiac surgery in infants remained relatively high, and it has been found recently in a large study of pediatric cardiac surgery that low weight at operation (<2.5 kg) was associated with mortality.[15]

The second finding also seems to be reasonable. Figure 5 shows the scatter plot of gestational week and birth weight for all of the patients in our series with EA. The white squares indicate patients who survived, and the black triangles indicate patients who died. The line of 1,500 g is far from the center of the birth weight distribution of infants with EA. Spitz et al[2] reported that the likelihood ratio for birth weight peaked at 1,500 g, but this finding is not sufficient reason to choose the value of 1,500 g as the cutoff value unless EA infants weighing <1,500 g were to have a clearly greater mortality rate than EA infants weighing >1,500 g. In the present study, the peak likelihood ratio for birth weight was at about 1,750 g, but this finding was not significant because the distributions of patients with EA according to survival were not separated clearly by birth weight, as shown in the scatter plot.

The discriminatory ability of our revised classification was verified using ROC analysis and found to be superior to that of the Spitz classification. The revised classification had good responsiveness for 2 different time periods, which suggests that the performance of this classification would be robust.

We evaluated the relationships among cause of death, associated anomalies, and the classifications. Our findings (Table V) seem to suggest the clinical problems of each of the revised classes. Patients in our revised class II had respiratory failure

Table V. Relation of classifications, anomalies, and causes of death

Spitz class	*Revised class*	*No.*	*Cardiac anomalies*	*Major associated abnormalities*	*Causes of death*
I	II	1		Duodenal atresia	TEF recurrence, empyema thoracis
		2		BFPM	RF
		3		MAS, PPHN, anorectal anomaly	RF
II	II	4		Trisomy 18	Intra-abdominal bleeding, CF
		5		Right bronchial agenesis Anorectal anomaly	RF
II	III	6	PDA + ASD	Pure EA	Severe pneumonia, RF
		7	PA + ASD + VSD	Multiple anomalies	Treatment withdrawn
		8	HLHS		CF
		9	DORV	Trisomy 21	Severe pneumonia, RF
		10	DORV		Tension pneumothorax
		11	CoA	Pure EA	Intestinal graft necrosis, sepsis
		12	ACC	CDH + left bronchial agenesis	RF caused by left bronchial stenosis
II	IV	13	ASD + VSD	Multiple anomalies	Treatment withdrawn
		14	ACC		CF, TEF recurrence, empyema thoracis
		15	TA + ACC		CF
		16	PDA + VSD	Trisomy 18	CF
		17	PDA + VSD	Trisomy 18	CF
III	IV	18	PDA + VSD		CF
		19	PDA + VSD	Trisomy 18	CF
		20	DORV	Trisomy 18, cerebellar hypoplasia	Treatment withdrawn

ACC, Aortic-coarctation complex; *ASD*, atrial septal defect; *BPFM*, bronchopulmonary foregut malformations; *CF*, cardiac failure; *CoA*, coarctation of the aorta; *DORV*, double outlet right ventricle; *HLHS*, hypoplastic left heart syndrome; *MAS*, meconium aspiration syndrome; *PA*, pulmonary atresia; *PDA*, patent ductus arteriosus; *PPHN*, persistent pulmonary hypertension; *RF*, respiratory failure; *TA*, tricuspid atresia; *TEF*, tracheoesophageal fistula; *VSD*, ventricular septal defect.

caused by congenital or acquired respiratory problems, which may be associated with prematurity. Patients in our revised class III had relatively severe, associated cardiac anomalies, and the treatment for the cardiac anomalies, EA, and other major anomalies would be difficult, thus raising the risk of total care. Causes of death of case 10 and 11 (tension pneumothorax and intestinal graft necrosis) were not related to either the cardiac anomalies or low birth weight directly, but rather reflect the difficulty of total care of EA patients in this class. Patients in our revised class IV had circulatory instability or cardiac failure caused by not only the cardiac anomalies, but also by prematurity. In this group of patients, trisomy 18 was a substantive problem with respect to survival. The new classification seems to reflect closely the clinical condition of EA infants, whereas the Spitz classification tends to include infants with various conditions of differing severity in the same class, especially in class II. With advances in neonatal intensive care and pediatric cardiology, the survival rate of our revised class II and III infants increased between the two 13-year time periods, and the survival rate of both classes became similar in the more recent time period. Nevertheless, classes II and III should be considered separately, because patients in these classes have different clinical conditions or problems, as shown in Table V.

Table VI. Responsiveness of the Spitz classification in reported EA series

Study (time period)	*N*	*AUC ROC (95% CI)*
Spitz et al[2] (1980–1992)	372	0.861 (0.79–0.93)
Okada et al[5] (1981–1995)	112	0.775 (0.65–0.90)
Choudhury et al[6] (1978–1997)	240	0.657 (0.56–.076)
Yagyu et al[7] (1960–1997)	133	0.850 (0.75–0.96)
Driver et al[8] (1986–1997)	134	0.667 (0.53–0.81)
Konkin et al[9] (1984–2000)	144	0.847 (0.74–0.96)
Lopez et al[4] (1993–2004)	188	0.822 (0.71–0.94)
Lilja and Wester[10] (1986–2005)	147	0.752 (0.62–0.89)
Present study (1980–2005)	121	0.795 (0.69–0.90)
Total (1960–2005)	1,591	0.747 (0.71–0.79)

The presence of major cardiac anomalies and low birth weight are valid predictors of survival in infants with EA. From the perspective of the Spitz classification, it is important to increase the survival rates of

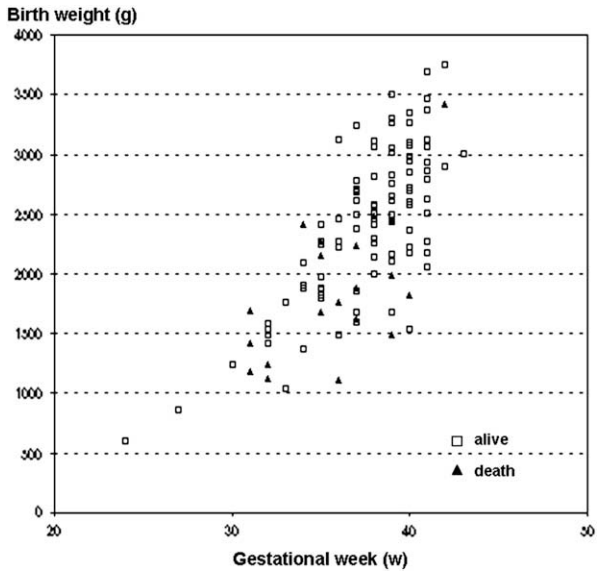

Fig 5. Scatter plot of gestational week and weight at birth.

EA babies with major cardiac anomalies or very low birth weight, and this continues to be an important challenge in modern pediatric surgery. Our revised classification seems to more accurately represent the clinical conditions of infants with EA and can be used to give the parents a realistic prognosis, as well as to compare results. Our classification also delineates the very low-risk patients (class I) with conventional surgical repair of EA. This information would be useful to compare results of new surgical procedure—for example, between laparoscopic EA airs and a conventional approach. In the future, it may be necessary to refine the definition of "major cardiac anomalies or circulation state" in babies with EA to better refine this classification. A prospective, multicenter study would, however, be needed to confirm the usefulness of this revised classification to increase the numbers of patients in each classes and to confirm our classification in other patients populations.

REFERENCES

1. Waterston DJ, Bonham Carter RE, Aberdeen E. Esophageal atresia: Tracheo-esophageal fistula. A study of survival in 218 infants. Lancet 1962:819-22.
2. Spitz L, Kiely EM, Morecroft JA, et al. Oesophageal atresia: At-risk groups for the 1990s. J Pediatr Surg 1994;29:723-5.
3. Harmon CM, Coran AG. Congenital anomalies of the esophagus. In: Grosfeld JL, O'Neill JA, Fonlkalsrud EW, Coran AG, editors. Pediatric surgery. Philadelphia: Mosby-Elsevier; 2006. p. 1051-981.
4. Lopez PJ, Keys C, Pierro A, et al. Oesophageal atresia: improved outcome in high-risk groups? J Pediatr Surg 2006; 41:331-4.
5. Okada A, Usui N, Inoue M, et al. Esophageal atresia in Osaka: a review of 39 years experience. J Pediatr Surg 1997;32:1570-4.
6. Choudhury S, Aschcraft KW, Sharp RJ, et al. Survival of patients with esophageal atresia: influence of birth weight, cardiac anomaly, and late respiratory complications. J Pediatr Surg 1999;34:70-3; discussion 74.
7. Yagyu M, Gitter H, Richter B, et al. Esophageal atresia in Bremen, Germany—evaluation of preoperative risk classification in esophageal atresia. J Pediatr Surg 2000;35:584-7.
8. Driver CP, Shankar KR, Jones MO, et al. Phenotypic presentation and outcome of esophageal atresia in the era of Spitz classification. J Pediatr Surg 2001;36:1419-21.
9. Konkin DE, O'Hali WA, Webber EM, et al. Outcomes in esophageal atresia and tracheoesophageal fistula. J Pediatr Surg 2003;38:1726-9.
10. Lilja EH, Wester T. Outcome in neonates with esophageal atresia treated over the last 20 years. Pediatr Surg Int 2008;24:531-6.
11. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148:839-43.
12. Poenaru D, Laberge JM, Neilson IR, et al. A new prognostic classification for esophageal atresia. Surgery 1993;113:426-32.
13. Teich S, Barton DP, Gin-Pease ME, et al. Prognostic classification for esophageal atresia and tracheoesophageal fistula: Waterston and Montreal. J Pediatr Surg 1997;32:1075-80.
14. Stevenson DK, Wright LL, Lemons JA, et al. Very low birth weight outcomes of the National Institute of Child Health and Human Development Neonatal Research Network, January 1993 through December 1994. Am J Obstet Gynecol 1998;179:1632-9.
15. Curzon CL, Miford-Beland S, Li JS, et al. Cardiac surgery in infants with low birth weight is associated with increased mortality: analysis of the Society of Thoracic Surgeons Congenital Heart Database. J Thorac Cardiovasc Surg 2008;135: 546-51.