Value of the International Classification of Diseases code for identifying children with biliary atresia

Pornthep Tanpowpong, MD, MPH, Chatmanee Lertudomphonwanit, MD, Pornpimon Phuapradit, MD, Suporn Treepongkaruna, MD

Division of Gastroenterology, Department of Pediatrics, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand

Background: Although identifying cases in large administrative databases may aid future research studies, previous reports demonstrated that the use of the International Classification of Diseases, Tenth Revision (ICD-10) code alone for diagnosis leads to disease misclassification.

Purpose: We aimed to assess the value of the ICD-10 diagnostic code for identifying potential children with biliary atresia.

Methods: Patients aged <18 years assigned the ICD-10 code of biliary atresia (Q44.2) between January 1996 and December 2016 at a quaternary care teaching hospital were identified. We also reviewed patients with other diagnoses of code-defined cirrhosis to identify more potential cases of biliary atresia. A proposed diagnostic algorithm was used to define ICD-10 code accuracy, sensitivity, and specificity.

Results: We reviewed the medical records of 155 patients with ICD-10 code Q44.2 and 69 patients with other codes for biliary cirrhosis (K74.4, K74.5, K74.6). The accuracy for identifying definite/probable/possible biliary atresia cases was 80%, while the sensitivity was 88% (95% confidence interval [CI], 82%–93%). Three independent predictors were associated with algorithm-defined definite/probable/possible cases of biliary atresia: ICD-10 code Q44.2 (odds ratio [OR], 2.90; 95% CI, 1.09–7.71), history of pale stool (OR, 2.78; 95% CI, 1.18–6.60), and a presumed diagnosis of biliary atresia prior to referral to our hospital (OR, 17.49; 95% CI, 7.01–43.64). A significant interaction was noted between ICD-10 code Q44.2 and a history of pale stool (*P*<0.05). The area under the curve was 0.87 (95% CI, 0.84–0.89).

Conclusion: ICD-10 code Q44.2 has an acceptable value for diagnosing biliary atresia. Incorporating clinical data improves the case identification. The use of this proposed diagnostic algorithm to examine data from administrative databases may facilitate appropriate health care allocation and aid future research investigations.

Key words: Administrative, Child, Cirrhosis, Coding, Jaundice

Key message

Question: What is the value of the diagnostic code in identifying cases of biliary atresia in a large administrative database?

Finding: The diagnostic code's accuracy and sensitivity are acceptable for identifying algorithm-defined cases. A history of pale stool and a presumed diagnosis of biliary atresia prior to referral added value.

Meaning: The addition of clinical data to the diagnostic code significantly increased the diagnostic yield.

Introduction

Biliary atresia (BA) is a progressive and fibrosing cholangiopathy that involves bile ducts which can lead to liver fibrosis and cirrhosis.[1] The incidence ranges from 0.3–3.7 per 10,000 live births,[2] but the aetiology remains unknown.[3] Initial symptoms include jaundice, pale stool, dark urine, hepatosplenomegaly, and ascites. Intraoperative cholangiogram is the currently accepted 'gold-standard' diagnostic investigation. Other investigations such as abdominal ultrasonography, hepatobiliary scintigraphy, and liver biopsy have been used to aid in the diagnostic process.[4,5] Once the diagnosis is made, hepatic portoenterostomy should be performed to restore adequate bile flow from the remaining patent biliary tract to the small intestine. However, most cases with failed hepatic portoenterostomy require liver transplantation by the age of 2 years due to cirrhosis and end stage liver disease; otherwise these children would die or suffer from severe cirrhosis-related complications throughout their lives. Most of the liver transplant candidates require on time referrals to a tertiary or quaternary care center specialized in pediatric transplantation with multidisciplinary team approach. Case identification for this devastating condition in the large administrative databases in conjunction with additional key clinical data may therefore help in health care allocation and prioritization.
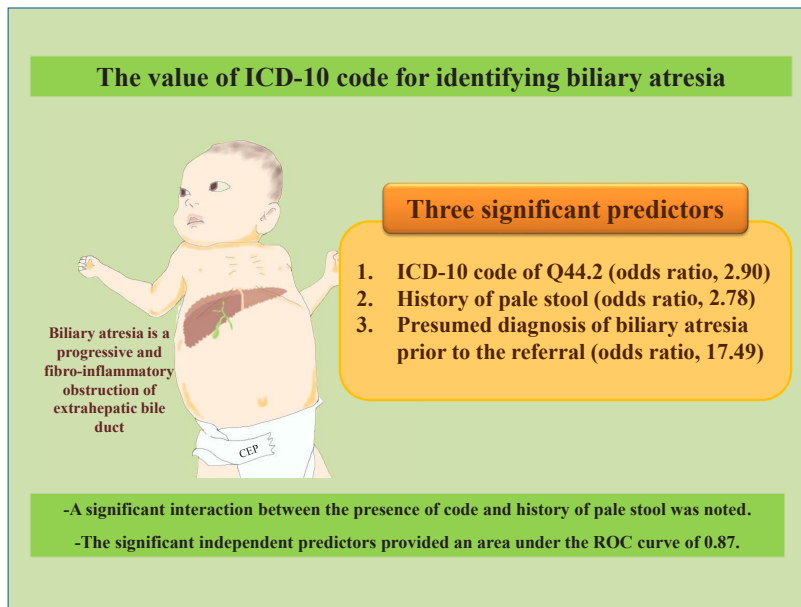
Administrative databases have been used for case identification in several diseases in children and adults, both in prevalent[6-8]

Graphical abstract. ICD, International Classification of Diseases; ROC, receiver operating characteristic.

and other less common diseases that pose significant burden to patients and the health care system.[9] The International Classification of Diseases, Tenth Revision (ICD-10) coding system has been implemented in various countries. The system provides a coding system based on the etiology and organ-based conditions or diseases. However, the primary purposes of this coding system aim for appropriate reimbursement and resource allocation in the healthcare system. Various underlying diseases, health care-related complications, or even health care settings may alter the coding assignment.[10] Several studies confirmed that using the ICD code alone would lead to a misclassification of various diseases, e.g., *Clostridioides difficile* infection,[7] food allergy,[8] and celiac disease.[11] Adding further clinical and laboratory information from the medical record has been demonstrated to improve the values of ICD code.[9,11,12] Jancelewicz et al.[13] used several clinical data points to create a diagnostic algorithm to help excluding patients with BA from other causes of neonatal cholestasis. These variables were liver chemistries, scintigraphy, and cholangiography; but data from the ICD code was not included. To our knowledge, no study has defined the value of ICD-10 diagnosis code for case identification of children with BA. We hope that this study can aid in identifying potential BA patients for future research studies and health care allocation; and hopefully with some key clinical data, the diagnosis code may be helpful in demonstrating the true prevalence of BA.

Methods

We created a cohort of all possible pediatric cases (<18 years) by performing an electronic administrative database search at a single center between January 1996 and December 2016. The hospital is a quaternary care, teaching hospital which is also one of the few active medical centers performing pediatric liver transplantation in the country. Patients were initially identified by screening health care organization databases and hospital billing coding system for the ICD-10 diagnosis code of Q44.2 for BA. Additionally, we also reviewed cases with secondary biliary cirrhosis (K 74.4), biliary cirrhosis, unspecified (K 74.5), and other and unspecified cirrhosis of liver (K 74.6) to potentially identify more patients with BA. The first documentation of an ICD-10 diagnosis code from the individual patient's record was considered as the initial encounter date. Cases without information on the initial presentation were excluded.

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee on 3 August 2013 (The Committee on Human Rights Related to Research Involving Human Subjects, Faculty of Medicine Ramathibodi Hospital, Mahidol University - ID 03-56-76) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. As this is a retrospective medical record review, the patient consent form was waived.

1. Patient and public involvement

As this is a retrospective study, the patients were not involved in the development, design, recruitment, conduct, reporting or dissemination of the study.

2. Medical record review for potential variables associated with diagnosis of BA

In conjunction with the ICD-10 diagnosis code, we randomly sampled patients for an electronic medical record review. We gathered and entered information in the structured data abstraction form. Information regarding demographic data, clinical characteristics, investigations, and operations performed
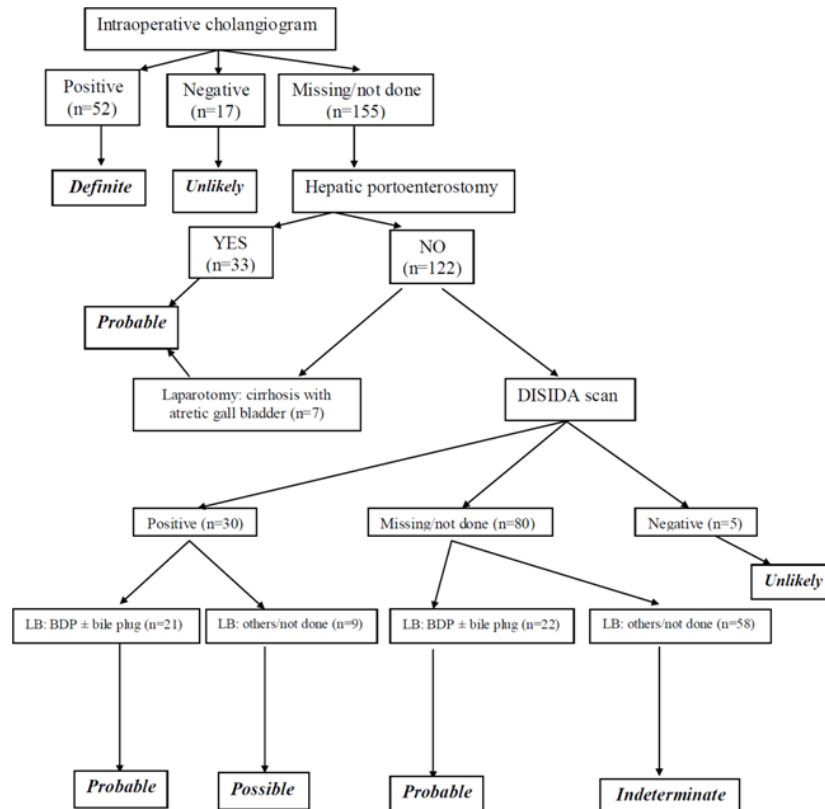
Fig. 1. Diagnostic algorithm for identifying children with biliary atresia. LB, liver biopsy; BDP, bile duct proliferation.

prior to the referral and at our institution, diagnosis prior to the referral and the final diagnosis were included in the study. The data abstraction form was initially verified by gastroenterologists and cross-checked prior to the formal medical record review.

3. Diagnostic algorithm for BA

To have a unified assessment of BA, we created a diagnostic algorithm that applies key information including hepatobiliary scintigraphy, liver biopsy, intraoperative cholangiogram, and history of hepatic portoenterostomy (Fig. 1). Clinical data were not initially included in the algorithm because of their suboptimal yield in the diagnosis of BA,[1] but the data may later be included in the further univariate and multivariate analyses. The algorithm classified cases into 5 categories: definite, probable, possible, indeterminate, and unlikely. The algorithm was also critically reviewed by experience pediatric gastroenterologists in the field.

4. Biostatistical analyses

We performed all analyses using Stata ver. 15.0 (StataCorp LP, College Station, TX, USA). Data are expressed as mean (standard deviation, SD), median (interquartile range, IQR), and proportions (with 95% confidence interval [95% CI]). Comparisons of discrete variables across different groups were assessed using chi-square test or Fisher exact test, while we used student t test or analysis of variance test for continuous variables. We calculated the overall positive predictive value (PPV) of the ICD-10 code for identifying "definite," "probable," or

"possible" BA cases with all patients with ICD-10 Q44.2. Then, logistic regression was used to examine unadjusted associations between variables of interest and definite/probable/possible BA. The multivariable models were created via forward stepwise selection (using $P<0.1$). The multivariable models were tested with Hosmer-Lemeshow goodness-of-fit tests. We determined the performance of the models by an area under the receiver operating characteristic (ROC) curve. An area under the ROC curve>0.8 is considered excellent, and 0.5 indicates that the predictive model is likely due to chance alone. A 2-tailed $P<0.05$ was considered statistically significant.

Results

We reviewed 162 patients with ICD-10 code Q44.2 (i.e., coded BA) and 76 patients with ICD-10 codes K74.4, K74.5, or K74.6 (i.e., coded non-BA) during the study period. We could not find information on the initial presentation from 14 medical records (7 in each group). Therefore, we finally analyzed 224 children (155 in the coded BA and 69 in the coded non-BA group). Table 1 represents baseline characteristics, ICD-10 information, and medical history. Information regarding abdominal ultrasonography and liver biopsy were limited (n= 144 and 110, respectively [data not shown]). Most patients in both groups were referrals from other hospitals (91% and 85%, respectively), while 86% of the coded BA group had a presumed

diagnosis of BA prior to the referral, but only 27% in the coded non-BA group.

By applying the proposed diagnostic algorithm for identifying potential cases of BA, we determined "definite" in 23%, "probable" in 37%, "possible" in 4%, "indeterminate" in 26%, and "unlikely" in 10%. Some patients with "probable" BA had a hepatic portoenterostomy performed but did not have information regarding intraoperative cholangiogram. Furthermore, 7 patients had documented cirrhosis and atretic gall bladder on exploratory laparotomy, so the decision was made to defer the hepatic portoenterostomy even the diagnosis was likely BA; therefore we proposed "probable" BA for these patients as well.

Overall, we found that the accuracy for identifying definite/probable/possible cases with BA was 80% with a sensitivity of 88% (95% CI, 82%–93%) and PPV of 81% (95% CI, 74%–87%). Further analyses revealed significant unadjusted predictors for definite/probable/possible BA were the following: age at initial presentation to our institution, presence of ICD-10 Q44.2, history of pale stool, a presumed diagnosis of BA prior to the referral to our center (Table 2).

In the final multivariable model (Table 3); 3 variables remained independent predictors of definite/probable/possible BA: the presence of code, history of pale stool, and a presumed diagnosis of BA prior to the referral. A significant interaction was noted between the presence of code and history of pale stool. The model passed the Hosmer-Lemeshow goodness-of-fit test ($P=0.51$) and could differentiate patients with vs. without definite/probable/possible BA with an estimated accuracy of 87% and an area under the ROC curve of 0.87 (95% CI, 0.84–0.89) indicating an excellent predictive capacity. After stratifying by the presence/absence of code Q44.2, if both the code and history of pale stool were not present, 28 of 29 children (97%) were noted to be in the "indeterminate/unlikely" group for BA diagnosis.

Table 1. Children's baseline characteristics and medical history by the ICD-10 code

Characteristic	ICD-10 Q44.2 (n=155)	ICD-10 related to cirrhosis of liver (n=69)
Demographics		
Age at the first encounter at our institution (mo), median (IQR)	7 (4–12)	19 (6–59)
Female sex (%)	59	55
Clinical presentations (available n=208)		
Pale stool (%)	72	48
Dark urine (%)	30	18
Steatorrhea (%)	12	7
Abdominal distention (%)	20	23
Hepatomegaly (%)	60	50
Splenomegaly (%)	47	38
Birth history (available n=187)		
Gestational age > 37 weeks (%)	90	93
Birth weight > 2,500 g (%)	84	83
Presumed diagnosis prior to referrala)		
Biliary atresia (%)	86	27

ICD-10, International Classification of Diseases, Tenth Revision; IQR, interquartile range.
a)204 of 224 [91%] patients were referrals.

Table 2. Unadjusted predictors of definite/probable/possible (vs. indeterminate/unlikely) biliary atresia

Characteristic	Odds ratio (95% CI)	P value
Age at the first encounter at our institution (every 1-year increase)	0.82 (0.74–0.90)	**<0.001**
Age (<6 mo vs. older)	1.92 (1.08–3.33)	**0.03**
Female (vs. male)	1.24 (0.71–2.16)	0.44
ICD-10 diagnosis code for atresia of the bile duct (Q44.2) vs. other cirrhosis (K74.4, K74.5, K74.6)	12.31 (6.29–24.11)	**<0.001**
Clinical presentations		
Pale stool	4.66 (2.49–8.70)	**<0.001**
Dark urine	1.40 (0.71–2.77)	0.33
Steatorrhea	1.32 (0.49–3.55)	0.58
Abdominal distention	0.77 (0.38–1.55)	0.41
Hepatomegaly	0.89 (0.49–1.61)	0.70
Splenomegaly	0.86 (0.48–1.54)	0.60
Presumed diagnosis prior to referral		
Biliary atresia vs. other diagnoses related to cirrhosis	35.28 (15.43–80.63)	**<0.001**

CI, confidence interval; ICD-10, International Classification of Diseases, Tenth Revision.
Boldface indicates a statistically significant difference with $P<0.05$.

Table 3. Adjusted predictors of definite/probable/possible (vs. indeterminate/unlikely) biliary atresia (available data, n=200)

Characteristic	Odds ratio (95% CI)	P value
ICD-10 diagnosis code for atresia of the bile duct (Q44.2) vs. other cirrhosis (K74.4, K74.5, K74.6)	2.90 (1.09–7.71)	**0.03**
Pale stool	2.78 (1.18–6.60)	**0.02**
Presumed diagnosis of biliary atresia vs. other diagnoses related to cirrhosis prior to the referral	17.49 (7.01–43.64)	**<0.001**

CI, confidence interval; ICD-10, International Classification of Diseases, Tenth Revision.
Boldface indicates a statistically significant difference with $P<0.05$.

Discussion

Our study examined the value of using ICD-10 diagnosis code of Q44.2 (i.e., a diagnosis code for BA) for identifying potential cases in the large hospital database. We found that the performance of the code was acceptable with an overall accuracy of 80% and PPV of 81%. The multivariable model by applying the presence of code, history of pale stool, and a presumed diagnosis of BA prior to the referral to our institution provided an excellent ability in defining definite/probable/possible cases with BA (area under the ROC curve=0.87). According to these findings, the aforementioned key information that can be obtained from the initial hospital visit in conjunction with the presence of code from the administrative database may be sufficient to define cases with BA. Previous studies also demonstrated that the ICD system alone did not perform well in some common gastrointestinal conditions,[7,11] and additional medical history from the medical record may be needed to correctly identify cases.[9,12]

Besides jaundice, pale stool may indicate extrahepatic biliary tract obstruction that is commonly found in BA. Dark urine, steatorrhea due to impaired bile flow causing fat maldigestion, hepatosplenomegaly can also be noted in infants with BA; however, these manifestations were not predictive in our study. BA usually requires further diagnostic investigations such as hepatobiliary scintigraphy, liver biopsy, or intraoperative cholangiogram. The significance of the presumed diagnosis of BA prior to the referral indicates that physicians were likely aware of this condition and proceeded with further investigations and management before referring these children to our institution.

Age at the initial visit to our institution was a significant predictor in the unadjusted model, but not in the multivariate model. Some patients may have persistent jaundice and pale stool after the hepatic portoenterostomy which indicate inadequate or poor bile flow to the bowel and lead to bile-induced hepatotoxicity, liver fibrosis, and eventually cirrhosis.[1] These patients usually develop signs of chronic liver disease and portal hypertension early, i.e., by 2 years old.[14] Referring physicians may decide to refer this group of patient with a presumed diagnosis of BA to our institution as a candidate for liver transplantation. Therefore, younger patients upon presenting to our institution were more likely to have a diagnosis of BA as compared to other non-BA patients (Table 2).

Our diagnostic algorithm may be applicable for identifying children with BA but may also have few limitations. For example, some patients had hepatic portoenterostomy performed without a strong evidence of BA, e.g., deferring intraoperative cholangiogram and/or liver biopsy. Previous reports showed that hepatic portoenterostomy was performed in some infants with Alagille syndrome[15] and other causes of cirrhosis that can potentially lead to a misclassification of disease. We were aware that liver biopsy provides a high sensitivity and specificity,[13,16] but the diagram was initially designed to aim for "excluding" extrahepatic biliary obstruction (e.g., BA) among patients with excreted hepatic scintigraphy (the "unlikely" group in Fig. 1) first before considering information on liver biopsy. Moreover, liver biopsy had just become more available at our institution during the recent years.

Although we acknowledge that a retrospective study from academic teaching hospitals may have incomplete information and be subject to potential recall and referral biases, we suggest that the findings from our study provide useful information on the value of using a large administrative database for case identification of BA. Performing multicenter studies to include various acuity settings may provide a useful tool for population at large. We also did not have the information on the ICD-10 code of each patient at the referral hospital. Gathering cases with BA can aid in the health care allocation and planning because this devastating condition leads to significant morbidity and mortality as well as the need for liver transplantation early in life.

Generally, BA is thought-out as a "single diagnosis code for one disease," but further investigations would be required to identify potential bias (e.g., coding errors) of using the ICD-10 code in various settings. We did not decide to include the diagnosis codes for neonatal jaundice or cholestasis because these codes would likely be too broad to identify potential cases with BA and may create a falsely high negative predictive value and specificity. We were aware that a sensitivity of ICD-10 diagnosis code (i.e., the proportion of patients who have the code among true cases) would also be helpful in determining the accuracy of diagnostic code. Once sensitivity and specificity are verified, the diagnosis code can become an important tool to obtain the actual prevalence of disease from large administrative national/international databases. We hope that our proposed algorithm can provide a platform for the future cohort creation.

In conclusion, we found that the performance of ICD-10 diagnosis code for BA provides a good value on sensitivity and PPV. Independent predictors of definite/probable/possible BA were the presence of ICD-10 code Q44.2, history of pale stool, and a presumed diagnosis of BA prior to the referral to our center. The final model including these 3 variables had an excellent predictive capacity for identifying children with BA. This strategy for case identification may be useful in facilitating future research studies and health care allocation. Moreover, studies using the large administrative database in conjunction with key clinical information may aid in demonstrating the true prevalence of BA.

84 Tanpowpong P, et al. Value of the ICD code in biliary atresia

www.e-cep.org

References

1. Hartley JL, Davenport M, Kelly DA. Biliary atresia. Lancet 2009;374:1704-13.
2. Jimenez-Rivera C, Jolin-Dahel KS, Fortinsky KJ, Gozdyra P, Benchimol EI. International incidence and outcomes of biliary atresia. J Pediatr Gastroenterol Nutr 2013;56:344-54.
3. Asai A, Miethke A, Bezerra JA. Pathogenesis of biliary atresia: defining biology to understand clinical phenotypes. Nat Rev Gastroenterol Hepatol 2015;12:342-52.
4. He JP, Hao Y, Wang XL, Yang XJ, Shao JF, Feng JX. Comparison of different noninvasive diagnostic methods for biliary atresia: a meta-analysis. World J Pediatr 2016;12:35-43.
5. Verkade HJ, Bezerra JA, Davenport M, Schreiber RA, Mieli-Vergani G, Hulscher JB, et al. Biliary atresia and other cholestatic childhood diseases: Advances and future challenges. J Hepatol 2016;65:631-42.
6. Lindenauer PK, Lagu T, Shieh MS, Pekow PS, Rothberg MB. Association of diagnostic coding with trends in hospitalizations and mortality of patients with pneumonia, 2003-2009. JAMA 2012;307:1405-13.
7. Chan M, Lim PL, Chow A, Win MK, Barkham TM. Surveillance for Clostridium difficile infection: ICD-9 coding has poor sensitivity compared to laboratory diagnosis in hospital patients, Singapore. PLoS One 2011;6:e15603.
8. Clark S, Gaeta TJ, Kamarthi GS, Camargo CA. ICD-9-CM coding of emergency department visits for food and insect sting allergy. Ann Epidemiol 2006;16:696-700.
9. Thirumurthi S, Chowdhury R, Richardson P, Abraham NS. Validation of ICD-9-CM diagnostic codes for inflammatory bowel disease among veterans. Dig Dis Sci 2010;55:2592-8.
10. Hsia DC, Krushat WM, Fagan AB, Tebbutt JA, Kusserow RP. Accuracy of diagnostic coding for Medicare patients under the prospective-payment system. N Engl J Med 1988;318:352-5.
11. Tanpowpong P, Broder-Fingert S, Obuch JC, Rahni DO, Katz AJ, Leffler DA, et al. Multicenter study on the value of ICD-9-CM codes for case identification of celiac disease. Ann Epidemiol 2013;23:136-42.
12. Thirumurthi S, Desilva R, Castillo DL, Richardson P, Abraham NS. Identification of Helicobacter pylori infected patients, using administrative data. Aliment Pharmacol Ther 2008;28:1309-16.
13. Jancelewicz T, Barmherzig R, Chung CT, Ling SC, Kamath BM, Ng VL, et al. A screening algorithm for the efficient exclusion of biliary atresia in infants with cholestatic jaundice. J Pediatr Surg 2015;50:363-70.
14. Chung PH, Wong KK, Tam PK. Predictors for failure after Kasai operation. J Pediatr Surg 2015;50:293-6.
15. Lee HP, Kang B, Choi SY, Lee S, Lee SK, Choe YH. Outcome of Alagille syndrome patients who had previously received Kasai operation during infancy: a single center study. Pediatr Gastroenterol Hepatol Nutr 2015;18:175-9.
16. Lee JY, Sullivan K, El Demellawy D, Nasr A. The value of preoperative liver biopsy in the diagnosis of extrahepatic biliary atresia: a systematic review and meta-analysis. J Pediatr Surg 2016;51:753-61.