The Causal Effect of Early Surgical Intervention on Biliary Atresia Patients Clinical Outcome Using Economic and Education Models

Eri Hoshino

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee of the Graduate School of Arts and Sciences
COLUMBIA UNIVERSITY
2015

ABSTRACT

The Causal Effect of Early Surgical Intervention on Biliary Atresia Patients Clinical Outcome Using Economic and Education Models

Eri Hoshino

Purpose and Background:

The aim of this paper is to investigate the causal effect of the timing of Kasai procedure, a surgical intervention, on Biliary Atresia (BA) patient outcomes using quasi-experimental techniques developed in economics and education. BA is a rare pediatric disease and is the most common cause of liver disease-related death in children. The causes of BA are not well established and are probably multifactorial. The Kasai procedure is usually performed at the early stage of disease, within the first few months of a patient's life. Although performance of the Kasai procedure before 60 days of age suggests improved prognosis, no available research to date has investigated the causal effect of the timing of the Kasai procedure on patient outcomes.

Methods:

This study uses data drawn from the Japan Biliary Atresia Registry. The dataset contains a total of 2743 patients; 1743 girls, and 1000 boys registered since 1989 to 2012. Subjects were eligible for study inclusion provided that they were born between 1989 and

2012 and underwent the Kasai procedure. In addition, eligible subjects had to have had an observable primary predictor variable and outcome. The primary predictor variable was the age at Kasai procedure and the outcome was 1-year native liver survival without jaundice. Other variables included in the analyses were sociodemographics (i.e. gender, birth order, birth weight, gestational age, parental age), clinical characteristics (i.e. types of obstruction, associated anomalies, type of choleretic agents), and hospital characteristics defined as hospital caseload. Bivariate analysis, multivariate logit/probit regressions, stratified multivariate logit regression, hospital fixed effects, and instrumental variable (IV) approaches were conducted.

Findings:

The IV approach was used to examine whether the timing of the Kasai procedure was endogenous in the model. Although using the length of stay at the last hospital prior to Kasai procedure as a proxy for accessibility to the nearest high-level hospital had its strengths as the instrument as confirmed by the first-stage diagnostics, the Wu-Hausman test indicated that the estimates of the model with and without IV were consistent.

The results from the multivariate covariate-adjusted logit regression suggested that for patients with obstruction types I and I with cyst, the timing of Kasai procedure was not associated with 1-year native liver survival.

However, in patients with type III obstruction (the majority of BA patients [86%]), when comparing to the timing of Kasai procedure at 61 to 75 days, the timing of Kasai procedure at less than 30 days was suggestively associated with the primary outcome of 1-year native liver survival without jaundice (log-odds ratio: 0.43 [t=1.79]). This was 1.5 times as likely to achieve 1-year native liver survival without jaundice while holding other variables constant. The timing of Kasai procedure after 76 days had a statistically negative association (log-odds ratio: -0.36 [p<0.05]) with the outcome. Moreover, the Kasai procedure at 106 to 121 days of age decreased the log odds of 1-year native liver survival without jaundice by 0.83, which equated to an increase in likelihood of achieving the primary outcome of 0.43 times. The Kasai procedure after 121 days of age also decreased the log odds of the outcome by 1.17 that was 0.3 times as likely to achieve 1-year native liver survival without jaundice. These differences are quite significant; therefore, the effort to increase the patients who could have the Kasai procedure at less than 30 days and to prevent the Kasai procedure after 76 days should be made to improve 1-year native liver survival without jaundice for the type III patients.

As previously suggested, cholangitis and the use of corticosteroid have a negative association with the outcome after adjusting for covariates. Ursodeoxycholic acid indicates a positive association with the outcome, as does the hospital caseload, even when both are

adjusted for covariates.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ACKNOWDGEMENTS

I would like to express my gratitude to Prof. Henry M. Levin for his support and patience through the course of my research. I would also like to thank Prof. Judith Scott-Clayton for her continued support and input. I could not have completed this research without their excellent guidance. In addition, I would like to thank the rest of my thesis committee: Prof. Tsang, Prof. Basch, and Prof. Reback for their insightful comments and encouragement.

My sincere thanks also go to the Japan Biliary Atresia Society for providing me with access to their data. I would also like to thank Prof. Nio and Dr. Sasaki for their precious support and advice. Without their cooperation, it would not have been possible to conduct this research.

Thank you to Dr. Urayama for his advice, encouragement, and stimulating discussions. Special thanks to Dr. Moody for enlightening this endeavor and not letting me to give up before its completion. I would also like to thank Ms. Muroga for her continual support to complete this process.

Last but not least, I would like to thank my family – my parents, my husband and our children – for supporting me throughout my work.

Chapter I

INTRODUCTION

1.1. Introduction

This study aims to investigate the causal effect of the timing of a surgical intervention, the Kasai procedure, on Biliary Atresia (BA) patient outcomes by applying the methods employed in economics and education. A randomized experiment provides the most persuasive strategy for answering causal questions regarding the impact of specific interventions. However, designing a randomized control trial (RCT) is not always the best approach for answering a causal question in the fields such as economics, education, and medicine because of factors such as cost or ethical constraints. Instead of performing a RCT, observational studies offer an alternative source for developing scientific evidence regarding the effectiveness of different treatments; however, a major challenge for observational studies is to overcome bias due to unobserved confounding.

In the face of limited economic resources, prudent allocation and prioritization of spending are vital. In fact, the idea of prioritized spending is universal across fields. Decisions for resource allocation should to be based, preferably, on scientific evidence of relative effectiveness. To evaluate the effectiveness of any given intervention is to estimate

the causal effect of one treatment compared to another. However, scientific evidence on treatment effectiveness for rare diseases is particularly lacking, and therefore continued research is typically needed.

Development of statistical methods to estimate the effectiveness of Educational programs has provided us with a better understanding of important patterns within available data, and the ability to test specific hypotheses in economics and education. This progression in the development of analytical methodologies has contributed to a major evolution in educational research in the U.S. (Murnane and Willett, 2011).

As much of the research conducted economics and education is based on real-world observational data, it is important to effectively address the issue of confounding. In fact, the impact of confounders on the estimation of a causal treatment effect can be mitigated by methods developed and used in economics and education. Researchers in economics of education often try to learn from experiments that occur naturally in order to overcome the intrinsic limitations of the observational data. Lessons learned in their experiences and methodologies developed during the course of their research could be applicable to any other fields where financial and ethical constraints may limit the application of RCTs. In this dissertation, I am applying statistical models that are often used in the fields of economics and education, but I am applying them in a healthcare research setting.

BA is a rare pediatric diseases and is the most common cause of liver disease-related death in children (Hartley, Davenport and Kelly, 2009). The causes of BA are not well established and are likely multifactorial (Danks et al., 1977; Dillon et al., 1994; Hartley, Davenport and Kelly, 2009). The Kasai procedure is a corrective surgery done at the early stage of disease, within the first few months of onset/diagnosis. Patient's age at Kasai procedure is one of the main prognostic factors in infants with BA (Chardot et al., 2013). Although a performance of the Kasai procedure before 60 days of age is associated with improved prognosis, to date, there are no studies in the available literature exploring the causal effect of early Kasai procedure on patient outcomes.

The general overview of BA and the timing of the Kasai procedure and patient outcomes will be discussed in more detail in Chapter 2. In the following subsection, I will explain the background of this study followed by the crucial period of child development, the research question of the study, and the overview of the organization of the study.

1.2. Background

1.2.1. Rare Disease and Chronic Health Condition in Children

There is no single definition widely accepted and internationally used for rare diseases. Some definitions rely on the number of people living with a disease while other

definitions include other factors, such as the existence of adequate treatments or the severity of the disease. In the United States, the Rare Diseases Act of 2002 defines rare disease strictly according to prevalence. Rare disease is defined as "any disease or condition that affects less than 200,000 people in the United States," or about 1 in 1,500 people. The definition of rare disease includes reference to treatment availability, a lack of resources, and severity of the disease. While the term orphan disease is used as a synonym for rare disease, the difference is that the term orphan drug is a legal definition in the U.S. and EU, and is often used in the context of the orphan drug movement which began in the U.S. This movement spawned the U.S. Orphan Drug Act, which is legislation designed to facilitate the development and commercialization of drugs to treat rare diseases.

A rare disease is defined by the European Union as one that affects less than 5 in 10,000 of the general population and there are between 6,000 and 8,000 known rare diseases which affect 6 % to 8 % of the population in the course of their lives in the EU (EURORDIS, 2005). The EURORDIS, a non-governmental patient-driven alliance of patient organizations representing 692 rare disease patient organizations in 63 countries, also reports that 1 in 17 people will be affected by a rare disease at some point in their lives. What this suggests is that collectively rare diseases are not rare. While individually these

conditions are uncommon, as a whole group they are a significant cause of chronic illness in both children and adults.

Moreover, 10-20 million children and adolescents in the U.S. have some form of chronic health condition (Pediatrics, 2014). The American Academy of Pediatrics defines "chronic" as a duration that lasts from three months to a lifetime. Based on these numbers, about 15 to 27% of all children in the U.S. live with a chronic condition[1]. Living with a chronic health condition may disrupt a child's normal activities, and requires numerous hospitalizations and/ or home health care and extensive medical care. 80% of rare diseases have identified genetic origins involving one or several genes or chromosomal abnormalities, and are present in between 3% and 4% of births. Other rare diseases are caused by infections (bacterial or viral), or allergies, or are due to degenerative, proliferative or teratogenic (chemicals, radiation, etc). Some rare diseases are also caused by a combination of genetic and environmental factors. However, for most rare diseases the etiological mechanisms are unknown (EURORDIS, 2005). The EURORDIS also suggests that despite this great diversity, rare diseases have some common characteristics, which are as follows:

- ✧ Rare diseases are severe to very severe, chronic, often degenerative and life-threatening;

[1] http://www.childstats.gov/americaschildren/tables/pop1.asp

- ✧ The onset of the disease occurs in childhood for 50% of rare diseases;
- ✧ The quality of life of rare diseases patients is often compromised by the lack or loss of autonomy;
- ✧ Rare diseases are highly painful in terms of the psychosocial burdens: the suffering of rare disease patients and their families are debilitated by psychological despair, the lack of therapeutic hope, and the absence of practical support for everyday living;
- ✧ Incurable diseases, mostly lack effective treatment. In some cases, symptoms could be treated to improve quality of life and life expectancy;
- ✧ Rare diseases are very difficult to manage: families encounter enormous difficulties in finding adequate treatment.

As mentioned, while individually these entities are uncommon, as a whole group they are an important cause of chronic illness in both children and adults and the as suggested, 1 in 17 people will be affected by a rare disease at some point in their lives (EURORDIS, 2005). The complexity and unknown mechanism of each rare disease, using an aggregate measure to cover the full range of patients with that disease, is very challenging. At the same time, if successful, it can contribute to tailor-made policies based on the scientific research to ease the difficulties faced by the patients. In general, people with a rare disease have not been registered in identified databases, and many rare diseases were summed up as “other disorders” as a consequence. By fully utilizing the Japanese Biliary Atresia Registry including 2743 subjects, this study suggests the advantage and possibility of generating scientific evidence for the disease with very low

incidence in the population.

1.3. Research Question

In this dissertation, I aim to answer the following research question:

Does the timing of the Kasai procedure relative to onset of BA have a causal impact on patient outcomes?

With regard to this research question, my concern is that there might be an endogeneity problem, since I cannot fully account for all confounders. For example, in the case of BA studies, the timing of Kasai procedure and outcomes are closely related to unobserved progression and the severity of disease. A detailed discussion of the endogeneity problem will be presented in Chapter 3.

The strength of this study is that by using the largest available population-based dataset, it allows to control for many confounders that could not be accounted for in prior studies. In this study, the primary predictor variable was the age at which a patient had the Kasai procedure, and the outcome was 1-year survival of native liver without jaundice. Other variables included socio-demographics (e.g. gender, birth order, birth weight, gestational age, parental age), clinical characteristics (e.g. type of obstruction, associated anomalies, and type of choleretic agents), and hospital characteristics (e.g. hospital caseload).

Yet, even after controlling for these variables, there may still encounter endogeneity bias requiring to consider using alternative statistical methods to correct for the unmeasured confounding.

1.4. Overview of the organization of the study

This study is organized and presented in six chapters. The introductory chapter explains the background and broader context of the study, the research question, and an overview of its contents and organization. Following to the introductory chapter, the second chapter reviews the existing literature, and is divided into two main sections. The first section will focus on previous clinical studies of BA and provides an overview of the disease followed by more specific topics including the impact of Kasai procedure timing on patient outcomes, as well as the types of statistical approaches used in previous analyses. The second section will focus on the methodological approach used in the analysis. In particular, I reviewed existing literature that employs quasi-experimental techniques in economics and education. The potential methodological approaches that were commonly used in the fields of economics and education but that could be utilized in this study are discussed and summarized in Chapter 2 as well. In chapter 3, I will discuss the foundation of causal inference and potential threats to causal inference, especially in observational studies. In

addition, I will outline and explore five approaches to mitigate such bias, including multivariate logit regression, difference-in-difference (DiD), propensity score matching (PSM), fixed effects (FE), and instrumental variable (IV). In chapter 4, I will describe the data and sample selection. Additionally, I will discuss the estimation strategy and focus on the feasibility of conducting the five approaches suggested in Chapter 3. Chapter 5 will cover the results of the bivariate analysis, multivariate logit/ probit regressions, hospital FE, and IV analysis, and will provide discussions of our empirical analyses. The last chapter covers the summary of findings, limitations of the study and suggestions for further research.

Chapter II

LITERATURE REVIEW

2.1. Introduction

The review of existing literature will be broken up into two main parts. The first part will focus on previous clinical studies on BA. Specifically this subsection will provide an overview of the disease as well as exploring literature followed by the effect of the timing of Kasai procedure on patient outcomes, the statistical approaches used in the previous BA analyses, and the implementation of stool color card screening.

The second part will present key methodological approaches developed and used in economics and education, focusing on studies that employ quasi-experimental techniques in these fields. Since the goal of this study is to examine a casual effect of the timing of Kasai procedure on patient outcomes using observational data, it is important to consider the application of advanced methodologies developed in economics and education for the analysis of BA. The strength of using quasi-experimental methods is that it allows for the estimation of a causal effect of a treatment variable when an experimental design is not feasible. The potential methodological approaches that are commonly used in the fields of economics and education, but that could be utilized in this study will be discussed and

summarized in the conclusion of this chapter.

2.2. Overview of Biliary Atresia

In this subsection, I will review the overview of BA, which covers pathogenesis, epidemiology, symptom, diagnosis, types of obstruction, the Kasai procedure, postoperative management, liver transplantation, and prognosis of BA.

The causes of BA are not well understood and are probably multifactorial. Genetic factors may play a role in some cases, but infectious, toxic, or immunologic mechanisms are likely involved (Danks et al. 1977; Hartley, Davenport et al., 2009; Tracy et al., 1994).

Although the overall incidence is low (about one in 10,000 to 20,000 live births), there are variations within specific regions (Matsui et al., 1994; Danks et al., 1977; McKierna et al., 2000). The worldwide incidence rates of BA are presented in Figure 1.

Seasonal and geographic variations shown in Figure 1 have been suggested, but not confirmed, to influence outcomes (Yoon et al., 1997; Chardot et al., 2013; Wada et al., 2007). In terms of socio-economic status of patient, one US study showed that black mothers were 2.5 times more likely to give birth to an affected child than were white mothers. No association was identified between the disorder and smoking, maternal age, education, alcohol use, folic acid intake, gravidity, parity, parental income, infant sex, preterm birth,

infant birth weight, or plurality (The et al., 2007); however, the results of this study were compromised by the small number of cases (n=62).

Most infants with BA are born at full term, have a normal birth weight, and initially thrive similar to healthy infants. The first sign of BA is jaundice; the onset of which may occur any time from birth up to eight weeks of age. If this sign goes unrecognized, it may lead to liver failure and eventually death in the first few years of life. BA may occur in isolation (70 percent), in association with lateralization anomalies such as situs inversus or asplenia[2] (10 to 15 percent), or with other congenital malformations (10 to 15 percent). Those with lateralization anomalies have a somewhat worse prognosis than those without anomalies (Shneider et al., 2006; Davenport, 2006; Schwarz et al., 2013).

The diagnosis of BA is made with a series of imaging and laboratory tests, together with a liver biopsy to exclude other causes of cholestasis; however, a definitive diagnosis of BA is only made after a cholangiogram is performed, which is usually done intraoperatively. If the diagnosis of BA is confirmed, a Kasai procedure is performed. Infants should be evaluated as rapidly as possible because the success of the surgical intervention diminishes progressively with increasing age at surgery (Chardot et al., 1999; Chardot et al., 2013; Serinet et al., 2009; Lien et al., 2011).

[2] See the classification of each category Table 1.

The anatomic pattern of BA is also identified at time of the Kasai procedure. In Japan, the anatomic pattern is divided into the four categories shown in Figure 2. The study done by Superina et al (2011) was to evaluate the impact of types of anatomic pattern and specific clinical factors on achieving successful surgical drainage, as defined by achieving a total bilirubin less than 2.0 mg/dL anytime within the first 3 months following surgery using logistic regression. They also examined the risk factors associated with transplant-free survival over the follow-up period using Cox models. Type II or III (atresia at the porta hepatis) was associated with a lower probability of achieving successful drainage compared to Type I (atresia of the common bile duct, or so called correctable atresia), although this difference did not achieve statistical significance (OR 0.29, p=0.086). In their study, Type I with cyst was not classified separately, and was instead included in Type I. In contrast to the weaker association of anatomical features with bile drainage, types of obstruction were associated with survival. Specifically, Type II or III (atresia at the porta hepatis) was associated with a greater risk of transplant or death compared to Ohi Type I (HR 2.03, p=0.030).

The Kasai procedure is a corrective surgery[3] done at the early stage of BA, usually within the first few months of life. During this surgery, the surgeon removes any

[3] See Figure 3

problem bile ducts outside the liver. The small intestine is then attached to the liver, providing a pathway that can allow bile to drain from the liver. The Kasai procedure is not a cure for BA, and it is not effective if the bile ducts inside the liver are damaged or missing. However, in many cases, the procedure can improve liver function in patients with BA for many years post intervention. In Western countries, short-term clearance of jaundice, which is considered as a prognostic factor for long-term outcomes, was reported in 50–60% of cases (Davenport et al. 2011), 30–40% of the patients can reach 10 years with their native liver (Stringer 2008). It is recommended that all infants with BA undergo a Kasai procedure[4]. Age at Kasai procedure is a main prognostic factor in infants with BA (Lien et al., 2011; Chardot et al., 1999; Chardot et al., 2013). Although a Kasai procedure before 60 days of age is associated with improved prognosis, no literature has focused on a causal effect of early Kasai procedure on patient outcomes. More detail review of the effect of the timing of Kasai procedure will be discussed in the next subsection.

Postoperative management following a Kasai procedure consists of interventions such as the use of choleretics and possible use of anti-inflammatory medications, nutritional rehabilitation, fat-soluble vitamin supplementation, prevention of cholangitis and

[4] It is recommended as Grade 1B on UpToDate which is an evidence-based, physician-authored clinical decision support resource which clinicians trust to make the right point-of-care decisions.. A Grade 1B recommendation is a strong recommendation, and applies to most patients. Clinicians should follow a strong recommendation unless a clear and compelling rationale for an alternative approach is present.

management of portal hypertension. The use of choleretics[5] such as ursodeoxycholic acid is standard practice in BA. Observational studies suggested a number of possible benefits of treatment with ursodeoxycholic acid, ranging from enhanced weight gain to reduced episodes of cholangitis and improved bile flow, but definitive evidence from randomized trials is lacking (Meyers et al., 2003).

Clinical evidence does not support routine use of steroids[6] (corticosteroids) in the treatment of BA. This was shown in a randomized placebo-controlled trial of steroid treatment in 140 infants with BA (Bezerra et al., 2014). Steroid therapy was given for 13 weeks and outcomes were measured at 6 and 24 months post Kasai procedure. No statistically significant benefit in bile drainage at six months post operation was observed in the infants treated with steroids as compared with the placebo group. In addition, no statistically significant improvement in survival with native liver at two years of age was observed in the treatment group. Moreover, the infants treated with steroids had significantly earlier onset of serious adverse events as compared with those given placebo.

Cholangitis is a common complication in patients with BA who have undergone the Kasai procedure. The incidence of cholangitis in these patients is between 40 and 90

[5] See Table 1 for the descriptions of types of choleretics
[6] See Table 1 for the description of steroid

percent, with the majority of patients experiencing at least one episode prior to two years of age (Luo and Zheng, 2008; Oh et al., 1995).

Even in the current development of the liver transplantation technique and post-operation management, at least 60 to 80 percent of patients with BA will eventually require a liver transplant. Liver transplantation is indicated for portal hypertension complications (recurrent variceal bleeding or intractable ascites), growth failure, and progressive liver dysfunction (Davenport, 2006; Nio et al., 2010; de Vries et al., 2011).

Although long-term prognosis for BA patients is variable, the complementary and sequential approach of the Kasai procedure and liver transplantation affords long-term survival with upwards of 90 percent of BA patients surviving into adulthood (Barbara E. et al, 2008; Chardot et al., 2013).

Although performance of the Kasai procedure clearly improves overall survival, the long-term prognosis is difficult to predict. The three important prognostic factors for surgical outcome are the types of obstruction of BA, younger age at the Kasai procedure and the expertise of the surgeon and care center at which the Kasai procedure is performed (Chardot et al., 1999; Chardot et al., 2013). Children born without atresia, categorized as type I obstruction, have the lowest risk of death or transplantation by two years of age (Nio et al., 2006; Superina et al., 2011).

Clinical outcomes after many complicated surgical procedures vary widely across hospitals and surgeons, and in a real-world setting, it is difficult to measure the respective skill-level of surgeons. Previous studies suggest that surgical success is dependent on the expertise of the center and surgeon (Chardot et al. 2013; Davenport et al. 2004; Nio et al. 2003). It appears that a center that performs at least five Kasai procedures per year has a better success rate, as measured by 5- or 10-year long-term survival with native liver (Chardot et al. 2013; Davenport et al. 2004; Nio et al. 2003). In my hypothesis, a surgeon's proficiency increases with experience as a result of their individually accumulated caseload. However, the hospital caseload would not necessarily account for each surgeon's experience and skills, but rather the characteristics of the hospital. Empirical data are lacking on the relationships between technical skill and postoperative outcomes.

2.2.1. Timing of Kasai Procedure on Patients' Outcomes

The early diagnosis and the timing of Kasai procedure could have different yet significant short to long-term implications. The definitive diagnosis of BA is made intraoperatively. This means that if the diagnosis of BA is confirmed, the Kasai procedure is performed. Although the Kasai procedure before 60 days of age suggests improved prognosis, this widely accepted perspective is based on an observed association, not a

causal relationship.

The basis for testing the relationship between the timing of a Kasai procedure and patient outcomes is that a relatively small shift in the timing of the procedure (e.g. a few weeks to a month) would eventually have a relatively large effect on patient outcomes such as a clearance of jaundice and native liver survival. If no clearance of jaundice is achieved through the Kasai procedure or when complications of biliary cirrhosis appear later on, liver transplantation is eventually needed. The later in life the transplantation is performed, the lower the associated morbidity and mortality rates.

The study performed by Lien et al. (2011) showed that the introduction of a nationwide stool color card was associated with a decrease in the median time to diagnosis from 47 to 43 days and an increase in the percentage of patients receiving their Kasai procedure within 60 days (from 60% before the program to 74% after implementation). This small change in time to the Kasai procedure was associated with an increase in 5-year, jaundice-free survival with native liver from 27% to 64%, and in overall survival from 56% to 89%. (Lien et al. (2011)

The severity of the liver damage at the time of Kasai procedure may be affected by many factors, both known and unknown, which could eventually affect the outcome. It may be that age is only a gross reflection of the cumulative damage to the liver thought the onset

of the disease, which for some patients may begin at the prenatal stage of development. Apart from age, other factors may also alter the tempo of progression to cirrhosis and irreversible damage. Previous studies have suggested that the age at surgery may not be independent of other variables such as inflammation, fibrosis, size of ductules in the excised specimen, hepatic stellate cell activation, poor nutritional status, and BA phenotype.

The systematic review by Jimenez-Rivera et al. (2013) investigated the population-based literature to assess international variation of BA incidence and outcomes and newborn screening. In particular, their focus was on the relationship between the timing of the Kasai procedure and patient outcomes, centralization of surgical Kasai procedure[7] and implementation of screening for BA as a national health system.

Their inclusion criteria for the review were that studies should be population-based and also include outcomes of interest, including overall survival, survival with native liver, and the timing of Kasai procedure. A total of 3,128 reports were identified through searches; 2,081 were retained for screening after excluding duplicate records. Of these, 149 articles were reviewed in full and 36 articles qualified for inclusion. An additional 4 studies were identified for inclusion from the reference lists of review articles and recent pertinent publications search, and therefore a total of 40 articles were included in the review.

[7] A center that performs at least five Kasai procedures per year has a better success rate, as measured by 5- or 10-year long-term survival with the native liver has been suggested (Chardot et al. 2013; Davenport et al. 2004; Nio et al. 2003).

According to the results of the review, there were 14 studies that analyzed the relationship between the timing of Kasai procedure and patient outcomes. The conclusions presented and a summary of the findings in table 2. All studies except for one (De Vries et al., 2012) demonstrated improved survival of native liver when surgery was performed at an earlier age.

Tiao et al. (Tiao et al., 2008) studied 327 BA patients born between January 1996 and December 2003 in Taiwan. Kaplan–Meier curves were used to assess patient survival while a log–rank test provided statistical comparisons between those who underwent a Kasai procedure before 60 days old and those who went at least 60 days. Furthermore, patient survival with native liver was analyzed using time from birth until either death or liver transplantation. An analysis of overall survival used time from birth to death. The 5-year survival rate with native liver was higher in those who underwent Kasai procedure before 60 days of age than those underwent the procedure after that time point (53.7% *vs* 36.7%, $p = 0.017$). However, the 5-year overall survival rate was not different between those who underwent a Kasai procedure before 60 days age and those who had the procedure on or after 60 days (69.5% *vs.* 63.6%, $p = 0.572$). This study also found that the 5-year overall survival rate during 1999–2003 was higher than that during 1996–1998 (74.8% *vs* 61.1%, $p = 0.014$). Though their results were based on two group comparisons,

there was a statistically significant difference between early and late Kasai procedures, with those having an early Kasai procedure demonstrating a higher 5-year survival rate with native liver.

A study by Lien et al. included 191 patients treated between 1990-2005 (Lien et al., 2011). BA patients were divided into three cohorts; those in cohort A were born before a stool card screening program[8] (1990-2000); those in cohort B were screened by the stool card regional screening program (2002-2003); and those in cohort C were screened by the stool card universal screening program (2004- 2005). The relative odds ratios were computed using logit regression. The Kaplan-Meier method and a log-rank test were also used to assess factors affecting survival.

The timing of the Kasai procedure at less than 60 days was 49.4% and 65.7% in cohorts A and B+C, respectively (p= 0.02). The jaundice-free (total serum bilirubin <2.0 mg/dL) rate 3 months after surgery was 34.8% and 60.8% in cohorts A and B+C, respectively (p<0.001). The 3-year jaundice-free survival rate with native liver was 31.5% in cohort A and 56.9% in cohort B+C (p < 0.001), whereas the 3-year overall survival rates were 64.0% and 89.2%, respectively (p<0.001). The 5-year jaundice-free survival rate with native liver was 27.3% in cohort A and 64.3% in cohort B (p < 0.001), and the 5-year overall

[8] The detail of stool color card will be discussed in the next subsection.

survival rates were 55.7% and 89.3%, respectively ($p < 0.001$). In their study, jaundice-free survival with native liver was considered as the quality outcome. Cohort B+C had higher rates of 3- and 5-year jaundice-free survival with native liver than cohort A (OR 2.87, $p = 0.001$, and OR 4.80, $p = 0.001$, respectively).

Patients who received a Kasai procedure before 60 days of age had better 3- and 5-year jaundice-free survival with native liver than patients who received an operation after 60 days of age (OR 3.25, $p < 0.001$ and OR 2.63, $p = 0.02$, respectively). Quality outcome was defined as jaundice-free survival with native liver. All survival time was calculated after the date of the Kasai procedure. In their logistic regression analysis given in Table 3, the age at Kasai procedure was statistically significant (OR 3.25, $p < 0.05$ for <60 days) while controlling for stool color card program which was also statistically significant (OR 2.87, $p < 0.001$), prophylactic antibiotics which was statistically significant as well (OR 3.03, $p < 0.01$), and sex of child which was not statistically significant (OR 1.25, $p =0.58$ for male).

This analysis is more reliable than others in terms of controlling for the possible confounding; however, they only account for the use of stool color card, gender, and use of prophylactic antibiotics. There could be other unobserved characteristics that might affect both the timing of the Kasai procedure and outcomes but not in the model. Moreover, the use of prophylactic antibiotics might be affected by selection bias, in which medical doctors

might have prescribed prophylactic antibiotics for patients whose background characteristics were different from those who were not prescribed e.g. worse or better condition.

The main purpose of the study done by Hsiao et al. (2008) was the same as Lien et al. (2011), focusing on the evaluation of a color card screening process. The study included 74 neonates born between January 2004 and December 2005 in Taiwan who were divided into 2 groups according to region of birth. The chi-square test and Fisher's exact test for categorical variables and the Student *t* test for continuous variables were used. BA infants who received an early Kasai procedure (prior to 60 days old) had a higher jaundice-free rate (72% versus 33.3%, $p = 0.002$) than those operated on after 60 days. However, these results should be interpreted with caution as the analysis did not account for any potential confounding.

A study in Canada was conducted that included 349 patients who were born and diagnosed with BA from 1985 to 2002 (Shneider et al., 2006). In order to control for time-variant factors, they analyzed them by dividing subjects into two groups: the first era, January 1, 1985 to December 31, 1995; and the second era, January 1, 1996 to December 31, 2002. Kaplan-Meier survival analysis was done with overall patient survival, defined as the period starting at birth and ending at death or last follow-up; liver transplantation survival

beginning at transplantation and ending at death or last follow-up; and native liver survival starting at birth and ending at death, transplantation, or last follow-up.

Univariate analysis was done using rank-sum and log-rank tests. The median patient age at the time of the initial referral to a tertiary care pediatric hospital center was 55 days (ranged 1 to 191 days). Four year survival rates with native liver were assessed by Kaplan-Meier survival curves for patients undergoing a Kasai procedure at age ≤30, 31 to 90, and >90 days showed that 49%, 36%, and 23% ($p < .0001$), respectively. This difference continued through 10 years. A total of 40% (6 out of 15) in the first era and 72% (13 out of18) in the second era (*p*-value was not reported) were referred at age >90 days. The 2- and 4-year post-Kasai procedure native liver survival rates were 47% and 35% for the first era and 46% and 39% for the second era (*p*-value was not reported).

This analysis suggested that there was a statistically significant lower probability of survival with a native liver at 4 years post operation compared to Kasai procedures done at age ≤30, 31 to 90, and >90 days. However a post hoc test to determine factors leading to the differences were not performed. They did not observe a significant era effect on the post-Kasai operation native liver survival rates although they did not control any other possible confounders. They did not observe a significant era effect on the post-Kasai operation native liver survival rates, yet it seemed to have the late referral problem which

contributes to diminish post-Kasai operation native liver survival rate.

A study in Brazil (Carvalho et al., 2010)[9] included 513 patients born and diagnosed as BA from 1982-2008 in six Brazilian medical centers. 76.4% of total patients underwent Kasai procedure of the mean age of 82.6 days (±32.8 SD). The 4-year post-Kasai procedure survival with native liver was 36.8%, inversely correlated with age at Kasai procedure (54, 33.3, 26.6% for ≤ 60, 61-90, > 90 days, respectively). The statistical test and significance were not reported.

A study in Switzerland (Wildhaber et al., 2008) included 48 patients with BA born in between January 1994 and December 2004. Survival rates were calculated with the Kaplan-Meier method, and prognostic factors were evaluated with the log-rank test. Forty-three children out of 48 underwent Kasai procedure. Median age at Kasai operation was 68 days (range, 30 – 126). The 4-year survival with native liver after Kasai operation was 37.4%. However, 4 -year survival with native liver was 75% in patients who underwent Kasai procedure before 46 days, 33% in patients operated on between 46 and 75 days, and 11% in patients operated on after 75 days (p = 0.02). The authors observed that survival with native liver after Kasai procedure was linked to the clearance of jaundice post-procedure. The study also found that 2- and 5-year survival with native liver was

9 Only the abstract was available in English.

92.3%±7.4% and 83.9%±10.4% in patients with successful Kasai [10]operation, 60%±21.9% and 60%±21.9% in patients with intermediate results following Kasai procedure, and 4.8%±4.6% and 0% in patients with failed Kasai procedure ($p < 0.01$). It should be noted that the study used a relatively small sample size..

A study in France (Chardot et al., 1999) included all patients with BA born in the years 1986 to 1996. The study included 472 patients. 5- and 10-year actuarial survivals with native liver rates were 32% and 27%, respectively. In the univariate analysis, better rates of survival with native liver are associated with the performance of Kasai procedure, and the age at Kasai procedure (performed prior to 45 days of age). (41% 31% for 5 year by less than 45 days and over 45 day, 41% 26% in 10 years respectively).

In their multivariate analysis to investigate the predictors survival with native liver, overall survival, and survival with liver transplantation, Chardot et al. controlled for the anatomical pattern of BA, the associated anomalies, and the center with large experience in the management of BA patients. The study results are shown in Table 4. The relative risk of survival with native liver for the patients with Kasai procedure after 45 days of age when compared to less than 45 days was 1.4 (CI:1.03-1.9, $p<0.01$). The results also suggest that patients with type III anatomic pattern have a higher risk of not achieving survival of native

[10] The detail definition of “successful”, “intermediate”, and “failed” Kasai procedure were not stated in their paper.

liver (RR 4.1, CI 2.1-8.0, $p<0.01$), as do patients with associated anomalies (RR 1.7, CI 1.1.-2.5, $p<0.05$). Also, patients treated at a hospital with less than five Kasai operations per year were indicated to have a higher risk of failure to achieve native liver survival compared to hospitals with more than 20 Kasai operations per year (RR 1.6, CI 1.3-2.1, $p<0.01$).

The cut-off timing of Kasai procedure differed from that used in the multivariate analysis by Lien et al (2011). Chardot et al. divided the timing of Kasai procedure at the 45 day instead of the 60 day. As the authors note, the 45 day time point was used since, in their univariate analysis, an earlier Kasai procedure was associated with a better prognosis for patients who had the procedure prior to 45 days of age (41% 31% for 5 year by less than 45 days and over 45 day, 41% 26% in 10 years respectively).

In a 1999 study, Chardot et al. used the same national data source described above, but investigated the timing of Kasai procedure by different categorical groups and with more subjects (Chardot et al., 1999). The authors classified patients into three groups: group 1 (n = 30), no contraindication to the Kasai procedure, but orientation to de novo transplantation; group 2 (n = 380), age at Kasai operation <90 days; and group 3 (n = 60), age at Kasai procedure >90 days. Survival with native liver, survival after liver transplantation, and overall survival were assessed using the Kaplan-Meier method, and

were compared using a log-rank test.

The results of the Chardot et al. (1999) suggested that five-year and10-year survival with native liver was 35% and 30% in group 2 and 25% and 22% in group 3 (p = .03). Five-year overall survival was 57%, 74%, and 55% in groups 1, 2, and 3, respectively (p = .003). Unlike their previous study, the authors did not control for any confounding in this study.

In a 2013 study, Chardot et al. continued their research on the long-term outcomes (Chardot et al., 2013). This study explored the same population-based data but over a longer period (1986 to 2009), with a sample population of 1107 patients. Patients were divided into 3 cohorts according to their years of birth: 1986–1996, 1997–2002, and 2003–2009. Survival rates were estimated according to the Kaplan-Meier method. Univariate analyses were performed using a log-rank test. Multivariate analyses, when appropriate, were performed using a Cox regression model (presented in Table 5). Study results showed a median age at Kasai operation of 59 days, unchanged over time. Twenty-year survival of native liver was 39%, 32%, 28%, and 19% after Kasai operation performed in the first, second, third months of life or thereafter, respectively (p= 0.0002).

While controlling for the types of obstruction and associated anomalies, the relative

risk of failure to achieve a survival of native liver increase as the timing of Kasai procedure advances, when compared to the group with Kasai procedure after 90 days (RR 0.54 CI 0.37-0.76 for less than 31 days; RR 0.59 CI 0.45-0.75 for 31-60 days; RR 0.75 CI 0.37-0.79 for 61-90 days; $p<0.01$). In accordance with their previous study, the anatomic pattern of type III showed a worse prognosis (RR 0.69 CI 0.54-0.87 for Type 3; RR 0.47 CI 0.32-0.71 for Type 2; RR 0.13 CI 0.04-0.4 for Type 1) which was statistically significant at 0.01%. Also, associated anomalies showed the worse prognosis (RR 0.59, CI 0.45-0.78, $p<0.01$).

Unlike their previous study, the authors did not include hospital caseload in their multivariate analysis. This reason for this was that by comparing three cohorts according to the years of birth, the first cohort showed better results for Kasai procedures performed in hospitals with the highest caseloads (20 cases per year); however, the Kasai procedure outcome results were not statistically different according to the hospital caseloads in the 2nd and 3rd cohorts.

Another France-based study used the same national data that Chardot et al (2009, 2013) employed; although, the study included patient data up to 2002 (Serinet et al., 2009). A total of 695 out of 743 included subjects underwent a Kasai procedure. The 2-, 5-, 10-, and 15-year survival rates with native liver were 57.1%, 37.9%, 32.4%, and 28.5%, respectively. Median age at Kasai operation was 60 days and was stable over the study

period. Whatever the follow-up (2, 5, 10,or 15 years), survival rates with native liver decreased when age at surgery increased (<30, 31–45, 46–60, 61–75, and 76–90 days) and the difference among those groups was statistically significant (p=0.0001), as shown in Figure 4. Kaplan-Meier analysis and the log-rank test were performed. They concluded that that the earlier the Kasai procedure, the better its results, especially in the subgroup of 59 patients operated on within the first month of life. Yet the outcome was worse in 9 of 92 patients who underwent a Kasai operation in the first month of life, suggesting that early diagnosis might have been related to a different pathogenesis of the disease, associated with a worse prognosis.

The findings of this series of French studies described above are informative, particularly those that controlled for confounding. In addition, the studies which assessed different Kasai procedure timing cut-offs (i.e. those other than the standard 60 days of age) provide useful insights into the role of this timing in overall outcomes. Also they investigated the different timing of the Kasai procedure rather than looking at the widely accepted notation of the Kasai procedure before 60 days of age.

A study in the Netherlands assessed liver status and health perception among Dutch patients who survived 20 years after therapy and investigated whether the rate of transplant-free survival increases with time (De Vries et al., 2011; De Vries et al., 2012). The

authors used a Dutch national database for BA which includes 104 patients born between 1977 and 1988. The study results demonstrated that the 20-year transplant-free survival rate increased from 20% (10 of 49) in the 1977 to 1982 cohort to 32% (18 of 55) in the 1983 to 1988 cohort (p = .03). Survival was significantly lower when patients underwent surgery after 75 days when compared with 60 to 75 days (11% ±6% vs 42% ±10%, respectively, p = .03), as shown in Figure 5. The authors used the Kaplan-Meier method and log-rank test for group comparisons. Unlike the other studies presented above, their findings on survival of native liver showed a somewhat mixed result. The 20-year survival rate was not significantly associated with the age at surgical correction. Interestingly, the best prognosis group had the Kasai procedure at 60 to 75 days followed by 45 to 60 days. Moreover, they reported that one fifth of the 20-year transplant free survivors had no signs of cirrhosis.

The analysis by Karrer et al. (Karreret al, 1990) used data from a registry including 904 children with BA drawn from more than 100 institutions. Patients included in the registry were born between 1967 and 1989 in the U.S. The 5-year actuarial survival was 62.5% (Kasai procedure <30 days), 43.6% (Kasai procedure 31-60 days), 39.5% (Kasai procedure 61-90 days), 28.6% (Kasai procedure 91-120 days), and 28.8% (Kasai procedure >120 days; p=0.023) Kaplan-Meier analysis and statistical comparisons using the Cox-Mantel test were conducted. Mean age at which the Kasai procedure was typically performed at present

has not changed significantly.

While previous literature suggested that the timing of the Kasai procedure was associated with the patient outcomes, only three studies tried to control for covariates in their respective models. The Taiwan-based study (Lien et al., 2011) controlled for gender, use of stool color card, and prophylactic antibiotics in their multivariate analysis. the age at Kasai procedure before 60 days was statistically significant.

In Chardot et al (1999)'s multivariate analysis, factors such as the anatomical pattern of BA, associated anomalies, and each center's level of experience with the management of BA patients were addressed. The Kasai procedure after 45 days of age as compared to less than 45 days increased the relative risk of failing to achieve a native liver survival.

The later study by Chardot et al (2013) also controlled for the types of obstruction and associated anomalies, the relative risk of failing to achieve a survival of native liver increased with increasingly later timing of the Kasai procedure defined as starting from 30 days of age added by 15 days.

In summation, although the association between the timing of Kasai procedure and survival of native liver has been suggested in many studies, the causal effect of Kasai procedure have not yet fully been investigated. The possible categorization for the timing of

Kasai procedure will be less than 30 days and added up by 15 days since the earlier intervention seemed to larger impact on the later outcomes in France study with the largest sample size in the existing study (Chardot et al 2009, 2013). They found that the earlier the Kasai procedure, the better the prognosis. Especially in the subgroup of 59 patients operated on in the first month of life, wherein the outcome was better than in the children operated on later in their lives.

2.2.2. Implementation of Stool Color Card

Since the previous studies present findings suggesting that the earlier age at Kasai procedure is associated with an increase in the long-term prognosis of patients, these results further support the need for effective BA screening. In fact, effective screening is the key to initiating a successful early intervention, since the diagnosis of BA is made with a series of exams including imaging, laboratory tests, and liver biopsy to exclude other causes of cholestasis, all of which could take a few days to a few weeks to collect and assess. Even after the various diagnostic tests, definitive diagnosis of BA can only be made intraoperatively. Once the diagnosis of BA is confirmed, the surgeon performs a Kasai procedure. Infants should be screened and evaluated as rapidly as possible because the surgical intervention's success depends on such a short period of time, between few weeks

to few months of age. Therefore, in this subsection, I will overview the evolution of the most widely used screening tool, the stool color card, and will look at the evaluation of the effectiveness of such a screening.

Japanese Studies

In Japan, beginning in April of 1987, the question asking "what color is your baby's stool?" with a footnote recommending "jaundice infants with alcholic or white stools need to see a pediatrician immediately" was included in a Mother and Child Health Handbook[11]. To assess the impact of this introduction, a study performed by Matsui (Matsui et al., 1995) compared the age at time of Kasai procedure in two groups, before-and-after the implementation. The first group included 83 patients born from January 1982 to March 1987 (before introducing the specific question in the Mother and Child Health Handbook) and the second group consisted of 93 patients born between May 1989 and April 1991. According to their results, the age at Kasai operation did not significantly differ between the two groups (the level of significance was not reported).

Following the report in 1994, Matsui et al. conducted a pilot study for mass-screening by using a stool color card including the pictures of normal and pale

[11] In Japan, all parents received a Mother and Child Health Handbook from a local municipality when they report their pregnancy. The booklet includes forms for prenatal checkups, a birth notification form, and information about available prenatal classes. The handbook also works as a "fact-sheet" recorded by as both medical institutions and parent. It contains the information on the mother's health, the progress of pregnancy and childbirth conditions, and also of the newborn child's condition until the child reaches school age, including the child's health, development and vaccination history.

pigmented stools in Tochigi Prefecture in Japan (Matsui et al,, 1995). They distributed the color card at maternity hospitals and mothers were asked to compare the stool color of infants with the referenced colors printed on the card. Doctors checked for infant's jaundice and the suggested stool color by a mother, and they reported cases needing further attention. They screened 17,641 infants from 1994 to 1995 and identified two out of three infants with BA during the period. Although this pilot study covered a large number of infants, since the incidence rate of BA was quite low, and the number of infants screened was not enough to perform any statistical tests. In a subsequent report by Matsui at a National Institute of Health workshop in the U.S., he reported that the specificity was 99.9% and the sensitivity was 67% (Sokol et al., 2007). According to their analysis, among 147,337 children the reported pale stool color at one month identified ten who were affected by BA, but five other children for whom there was no claim for pale stool colors were also found to have BA. The average days from birth at the Kasai operation was 53 for the screened group and 84 days in the non-screened group.

Muraji (2012) reported that alcholic stool was observed only in 50% of the infants before the BA diagnosis based on the Japanese Biliary Atresia Registry data. The author analyzed the effect of a color card by comparing patient data before- and after-implementation of the color card screening in 1997. The result showed that the median

age at referral increased from 53 days of age in the before-implementation group (n=18) to 70 days in after-implementation group (n=9) (Muraji, 2012). Two of nine patients were detected by color card and four patients presented with the color card at one month check-up, two cases were false negatives and the other two were missed. Among the remaining three patients, two cystic type patients showed good outcomes, while the other did not present the color card at the one month check-up and was detected at 95 days.

Muraji (2012) also reported results from an analysis using the Japanese Biliary Atresia Registry. Yellow stool was observed in only 55.2% of patients by the parents (1,333 of 2,417) and only 24% of parents claimed that there was abnormality in stool color in their BA infants.

Positive results of color card screening had been reported from the pilot study in Tochigi in contrast to the report from Ibaraki (Muraji, 2012; Matsui et al., 1996) . There were ten regions in Japan which implemented the stool color card screening prior to expanding it to a nation-wide screening scheme. However, the official reports on those pilot studies other than Tochigi and Ibaraki were not publically accessible. By contacting a respondent at regional or district level, I realized that there were great disparities in implementation of studies among different regions, which led to mixed results. Some regions offered educational lectures to medical staff as well, but some regions did not.

Without any scientific investigation on the effect of color card screening, a nationwide screening program with the color card has been implemented since April 2012. The stool color chart was included in the Mother and Child Health Handbook and asked a caregiver to see doctors if stool color of an infant fell into the suggested colors (#1-3 in Figure 6). The effect of the nationwide implementation has not been reported yet.

Taiwan Study

Prior to Japanese national implementation of the stool color card, Taiwan introduced a national stool color card for the early identification of babies with BA in 2004. As reported in the previous subsection, Lien et al (2011) employed a logistic regression model to estimate the effect of using a stool color card that had been implemented in Taiwan. The study compared the outcome of patients before versus after the launch of the infant stool color card screening program and concluded that the stool card screening program was effective in achieving early diagnosis of BA patients. In the Taiwan-based study, the sensitivity was 0.9714 and the specificity was 0.9995.

However, in their analysis, only gender, prophylactic antibiotics, and age at the Kasai procedure were controlled, and as previous literature suggested, other clinical and hospital characteristics might have affected the results. In addition to distributing a color card, educational lectures were provided to medical staff. Also to enhance better

interactions among medical staff and caregivers for screening BA, posters were distributed in all hospitals and clinics. Detailed orientations to support the national color card screening might have affected the outcome as well.

Other screening methods for early diagnosis are to measure blood, urine, and stool. The early measurement of blood included conjugated bile acids in dried blood spots using tandem mass spectrometry, serum conjugated bilirubin after birth, and serum Apo C-II and III proteins using surface-enhanced desorption/ionization time-of-flight mass spectrometry (SELDI-TOF- MS). The early measurement of urine included measuring urinary sulfated bile acid (USBA). The early measurement of stool included measuring fecal conjugated bilirubin by near-infrared reflectance spectroscopy. However, none have been put into daily practice extensively, due to both cost and technical complexity (Wildhaber, 2012).

2.3. Quasi-experimental Techniques used in Health and Education

In this subsection, I will review the methodological approaches that were utilized in previous studies in the field of economics and education aiming to examine the causal relationship between health and education. After the general overview of the statistical approach for such causal inference, I will focus particularly on the observational studies using a quasi-experimental methods used for examining the relationship between health

and education in general, and how researchers have tried to overcome latent endogeneity problems.

In the field of economics and education, randomized trials are not as common as in clinical research; However they are becoming more popular in education (Angrist, 2004). A leading randomized study from the field of economics and education is the Tennessee STAR experiment designed to estimate the causal effects of smaller classes in primary school. Results from the various studies on the Tennessee STAR experiment have shown a large effect to smaller classes (Finn and Achilles, 1999; Krueger and Whitmore, 2001). Even in the absence of a real experiment, one could try to look for well-controlled covariates and/or natural quasi-experiments. The quasi-experimental study of class size by (Angrist and Lavy, 1999) illustrates how we can leverage available observational data by designing studies that exhibit characteristics of an experimental design. As in Tennessee STAR, in accordance to other previous studies, they suggested a strong link between class size and achievement[12]. As used in Angrist and Lavy (1999), the IV approach has been widely used in economics and education in their causal estimate. In the following subsection, I will look at more detailed methodological approaches taken in selected studies aiming to investigate the relationship between health and education.

[12] The outcome was the test scores, the explanatory variable was class size, the IV was the indicator for whether total enrollment was "just above" a multiple of 40. Their Maimonides' rules states (roughly) that no class size should exceed forty, so that if enrollment (treated as exogenous) is "just below" 40, class sizes will be bigger, whereas if enrollment is "just above" 40, class sizes will be smaller.

In this subsection, I will examine the evidence from recent non-experimental and quasi-experimental studies to look at the effect of health status on education. Endogeneity between health and education can arise from unobserved factors that affect both health and education. For example, not only do schools and teachers vary in several ways requiring great effort to collect all the relevant data on those variables, but also some school and teacher characteristics (e.g., teachers' motivation and principal's managerial ability) are difficult to measure. Also one can have difficulty measuring a child's innate ability, although often IQ is used as a proxy. In addition, parents' preferences and favoritism are difficult to measure. Although, such parental preferences for a child's health and education are likely to be correlated. This leads to another important source of bias called reverse causality.

Many researchers use height-for-age as a measure for capturing general health status of a child in observational studies (Victora et al., 2008; Jones et al., 2008; Glewwe, 2005; Glewwe, Jacoby and King, 2001; Glewwe and Kremer, 2006; Alderman, Hoddinott and Kinsey, 2006). The indicators of "healthiness" used in observational studies tend to be less disease-specific, instead focusing on how to capture a broad picture of a child's health.

Not all observational studies using secondary data address the issue of endogeneity; therefore, conclusions depend on so-called "naïve" estimates such as results from ordinary least squares (OLS) estimations of the effect of health and nutrition status on

education outcomes[13]. After addressing the weakness of such studies, I will look at the studies addressing potential biases.

The strength of quasi-experiment studies is that the mechanisms are analogous to experimental studies that ideally would eliminate the heterogeneity in unobserved variables by creating a counterfactual situation. Although in reality, such a counterfactual situation is not likely to be created due to violations in the original planned assignment. Some of the reviewed studies employ an exogenous shock measure as their instrument. Since such an exogenous shock is used as the instrument for a regressor, in the reviewed literature, health status in the early age, it overcomes the exclusion restriction which previous studies have often violated. As a result, the use of such exogenous shock corrects many of the shortcomings of results obtained from non-experimental studies. The summary results were presented in Table 6.

Grantham-McGregor et al. (2007) reviewed on cross-sectional studies in developing countries and conclude that there was an association between malnutrition and cognitive educational outcomes measured in cognitive score, IQ test, and reasoning and arithmetic and also in school grade attainment of children. Unfortunately, in their article, the

13 Alderman and others (2001) indicate that child health is three times as important for enrollment than suggested by "naive estimates" that assume that child health is predetermined by household choices in the presence of unobserved factors

authors specifically state “we do not assess causality”(p.63) (Grantham-McGregor, 2007). Although with the limitation for a causal inference, summarized results in their study showed the overall trend of the effect of malnutrition and education outcomes.

Victora et al. (2008) performed an analysis using data from five long-standing prospective cohort studies from Brazil, Guatemala, India, the Philippines, and South Africa. Their analysis looked at how maternal and child undernutrition affected adult outcomes, including height, schooling, income or assets, offspring birthweight, body-mass index, glucose concentrations, and blood pressure (Victora et al., 2008). The status of child’s undernutrition was measured by birth weight, intrauterine growth restriction, weight, height, and body-mass index at two years according to the new WHO growth standards. Controlling for potential confounding variables such as age, parents’ educational level, and child’s socioeconomic status, the authors found that the strongest positive predictors of schooling was height-for-age and there was 0.5 increased years of schooling ($p<0.01$) with one unit increase in height-for-age.

Jones et al. (2008) examined the association between household socio-economic status (SES) at birth and stunting of growth between South Africa and the Philippines. The data were from two longitudinal birth cohorts, the Birth to Twenty study in South Africa and the Cebu Longitudinal Health and Nutrition Survey (CLHNS) in the Philippines. While the

final data size was reduced due to limited number of observation especially in South Africa (n 2293 total; reduced to 450: stunting at 1 year) and 401 (stunting at 2 years) and CLHNS infants (n 2513 total; 1820 (stunting at 1 year) and 1710 (stunting at 2 years). The Philippines' infants were significantly more likely to be stunted at 1 year (32.6 v. s.8.7%) and 2 years (48.9 v.s. 21.1%) compared with South African children (Jones et al., 2008).

The authors performed logistic regression analyses controlling for age and parity as covariates. The results showed that SES measures including maternal education, refrigerator and television ownership, and water and toilet facilities were significant predictors of stunting at one and two years of age in the Philippines' cohort (OR 0.61:0.53-0.70CI at age one and OR 0.51: 0.47-0.58 CI at age two) but not for those in South Africa.

While the above-stated studies mostly did not have all of variables which should be included in the analyses neither look at a causal relationship between child's unfavorable health status and education outcomes, there were some studies addressing endogeneity and seek for solutions to correct such a problem.

Glewwe and Jacoby (1995) carefully examined a causal effect of malnutrition on education outcomes using cross-sectional data in Ghana. They investigated the relationship between malnutrition (height-for-age at preschool age) and delayed enrollment and

(ultimate) grade attainment on 1757 Ghanaian children aged 6–15 years in 1988–1989. They tried two approaches to account for endogeneity. The first was to use the family FE model. The underlining idea was that child health varies within families, but parental tastes for child health and education outcomes do not vary within the family. In this setting, within family correlation of child health and education outcomes should not be caused by any such correlation in parental tastes. More specifically, the parental tastes for child health and child education outcomes. The second strategy they use was IV approach. They used distance to nearby hospitals and maternal height as instruments. The IV method can be used when analyzing variation across households and not within households since these instruments do not vary across children in the same family. Therefore, these two approaches complement each other.

The key assumption used the IV approach is that the distance to nearby medical facilities and the mother's height affect child schooling only through their impact on child health status. Although the authors recognized the limitations of the two approaches, there were few critical points to be argued. The family fixed-effect approach does not take the variation in child genetic healthiness or variation of reallocation of unobserved education inputs across different children within the family. Failing to take this variation into account will lead both to an underestimate and overestimate of the effect of child health on education.

For example, if parents see a relatively sick child compared to his/her siblings in family, they may put more resources to sicker child. Or, they might decide to do the opposite, investing less in such a child's education. Parents may treat children differently according to age, which is affected by family income cycle and parental experience with earlier children as well.

For the IV approach, there was the strong assumption[14] that the height of the mother and the distance to the nearest medical facility can affect child's education outcomes only through child health. The distance to the nearest medical facility can be correlated to many community characteristics that may also affect education decisions such as access to school, other public facilities, and therefore, also affect school quality. It may also be affected by the degree of urbanization; and therefore, access to urban services and urban influences, which ultimately affected child's education. Also the validity of using a mother's height as the instrument was doubtful if we see the mother's height as the health and nutrition status of her, then it could affect her productivity and it may result into household income as well. The mother's productivity would also affect how she allocates resources on her child's education. Although in their study, exogeneity of height-for-age was rejected but exogeneity of family income measured by per-capita expenditures was not rejected. When

[14] Other assumptions for the IV approach will be discussed in Chapter 3 and 4.

treated as exogenous, family income had a significant impact on delays only.

The authors found that strong negative impacts of child health measured through malnutrition on delayed enrollment using both the IV and FE estimators. With a decline in height-for-age of one, 0.634 month decrease in enrollment and the t-value is 2.49. Unexpectedly, they found no statistically significant evidence that child health increases school attainment as measured by grade completed.

Other studies that used a cross-sectional data are at a disadvantage, since the current health condition is greatly affected by the past health status. The result of measuring only current health status and estimate the impact on education outcome may result in positive and statistically significant results. In fact, the problem of endogeneity permeates not only in cross-sectional studies but also in longitudinal studies. In the next sub section, I will review recent studies that use longitudinal data.

Glewwe, Jacoby, and King (2001) used panel data from 2,192 households in the Philippines. Height-for-age was measured at age 8 and unlike the other studies focusing on the quantity of education outcome and education outcome was measured by test scores (Glewwe, Jacoby and King, 2001). They use two approaches as Glewwer and Jacoby (1995) used in their early analysis on Ghana. The use of family fixed effect removed family averages of innate academic ability and all school quality variables since the siblings in the

same household go to same school. They also used an IV approach to address remaining endogeneity and use the height of the older child at age 24 months.

The same arguments made for the Ghana study could be also applied here. It was intuitive to see the decisions of resources allocation on health investment and also on educational inputs could be influenced by differences in innate ability among siblings in the same family, which may lead to correlation between early childhood health investments and primary school test scores. However, the authors argued that health investments made until child's 24 months of age could not be correlated with innate child academic ability. They referred to psychology studies that conclude that parents cannot observe children's innate ability until the child is older than 24 months; therefore, it was assumed to be uncorrelated to health investment made by their parents. This was a very strong assumption to make but was needed in order to deliver the economic identification.

The authors also argued that the impacts of parental education inputs before the child reached primary school age and the impacts of current child health status on primary school academic scores were negligible so they could be dropped from that equation. On the condition that parents could not see the children's innate ability until the child was older than 24 months, they used the height of the older sibling by age 24 months as the instrument for height-for-age. Although they made strong assumptions to make their

estimation strategy more credible, the assumed conditions were unrealistic.

The most unrealistic assumption was that health investments made from conception through 24 months of age cannot be correlated with innate child academic ability because parents do not observe children's intelligence until after the child reached at least 24 months of age. Many parents believe that the education inputs such periods have strong impact on later cognitive outcomes. With limitations of the estimation strategy explained above, the authors concluded that children's health status measured by height at age 8 had strong impacts on their test scores. They found that a one standard deviation increase in height-for-age raises the achievement test score by 8.9 (t=2.8) points. They also show that such direct effect of nutrition on learning productivity per year of school was equivalent to spending about eight extra months in school.

Alderman et al (2001) used panel data collected for 800 households from 1986 to 1991 in Pakistan. They used height-for-age at 5 years old and education outcome was measured by the probability of being enrolled in school at age seven school enrollment. In order to correct the endogeneity problem of health status, they used food prices as an instrument, which captured the deviations in prices from long-term trends obtain unbiased estimates.

Their study had several potential limitations. First, they did not control for

household heterogeneity. Since the IV method can be used when analyzing variation across households not within household, the variation within family was remained as untreated. The use of food price shocks in the first time period as IVs for health status in that time period was theoretically valid only if price shock affected education outcome only through the health status of that time periods. However, since they included an initial wealth variable as covariates, household expenditures averaged over three years which included those on educational inputs, given price shock was more likely to affect the spending on education as well.

Moreover, the price shock in the first time period also affected parents' behavior on resource allocation on child's education in the second time period by changing savings. Therefore, the use of price shock in the first time period potentially violated the exclusion restriction. With several limitations, their results suggested that the effect of height was significantly and substantially greater for girls than for boys. The results were seven times larger for girls than for boys (0.51 with SE=0.14 for girls; 0.07 with SE=0.18 for boys).These results were equivalent to an improvement of 0.25 in height-for-age score raise, the probability of school enrollment for girls by 9 percent over the base case and for boys by 2 percent.

The Institute of Nutrition of Central America and Panama (INCAP) study (Pollitt et

al., 1993) focused on the relationship between a high protein diet and education outcomes. It was initiated in four Guatemalan villages in 1969, two of which were randomly selected to receive a porridge (*atole*) that was high in calories and protein while the other two villages received a drink (*fresco*) with less calories and no protein. Follow-up studies over the next two decades appear to show sizeable effects on later cognitive outcomes from providing the *atole* to mothers and young children.

Focusing on the effect of child's health on education outcomes, Maluccio et al (2006) evaluated the middle-long term impact of nutrition supplement intervention in Guatematra. The attrition rate was relatively low because they set certain criteria[15] for inclusion for the follow-up study, there were 1,090 individuals (46 percent of the original sample and 54 percent of those alive in 2007) from the original sample and 1,463 children of original sample members. The exposure period chosen by the authors was defined as the period from birth to 36 months of age, based on earlier research with the 1969–77 data. For education outcomes, they used grade attainment, reading comprehension and non-verbal skills.

Maluccio et al. employed the DiD[16] approach, the village fixed effect and the IV

15 These criteria are who 1) were interviewed successfully 2) were living in one of the original study villages 3) had a biological parent living in the area

16 However, in their estimation strategy, only one reduced-form equation is written, the detail use of difference-in-difference is not specified clearly.

(exposure to *atole* as the instrument) to account for potential biases. The village fixed effect to capture all fixed characteristics of four villages that might affect education related outcomes. It was crucial to include these because of the small number of villages in the experiment. The village fixed effect could capture persistent cultural differences and fixed differences in educational or economic opportunities that might have resulted in different educational investments across villages even without the interventions. Maluccio et al also controlled for other confounding variables including gender, birth year of a child, parents' age when child was born, parents' educational attainment, household wealth, and the distance to feeding center. In addition, the authors constructed community-level covariates that related to key education-related decisions in a child's development. Those were indicator variables of the availability in the village of a permanent (cement-block) structure for the primary school when the respondent was seven and 13 years old (supposedly the time children enter primary school and secondary school), and primary school student-teacher ratios when the respondent was seven and 13 years old. This accounted for the time-variant characteristics of community and school. Their results showed that female schooling increased by 1.17 grade (t=2.13). Both female and male reading comprehension were increased by 0.28 scores (t=2.52) which was equivalent to about one-quarter of a standard deviation. Also for both female and male non-verbal skills were increased by 0.24

scores (t=2.01) which was also equivalent to about one-quarter of a standard deviation.

In the INCAP study, sample attrition was a major concern in the 1988–1989 follow-up, as 40% of the original sample was not included in the final analysis. The attrition in the sample might have been associated with a number of initial conditions and characteristic of that individual or household that affected health status and education outcomes, which depends on whether attrition was random or non-random and whether sample size was large enough for power tests. In order to account for the attrition bias, sensitivity analyses, they used the results from census data which included the same individuals from the original study and covered 90% of the original sample. Using this village census-based sample, the estimated effects of atole on grades attained were remarkably similar to those based on HCS, including the exact same set of controls.

The study by Linnemayr and Alderman (Linnemayr and Alderman, 2011) used data from the two survey rounds (2004 and 2006) of the Senegal Nutrition enhancement program which was included to provide participants with vitamin A and deworming for children 6–59 months, iron for pregnant women, bednets, breastfeeding promotion, cooking workshops and implemented from 2004 to 2006. The authors used weight for age as the measure for child health status, which tends to indicate not only a child's health status in the long-term, but also the short-term. This could be problematic since this measure did not allow for the

isolation chronic and acute nutrition status, and possible factors that might have affected one and not the other. However, in this study, a serious selection bias existed. The result showed that of the 111 villages that were initially assigned treatment status, 80 ended up receiving treatment, while 31 villages (28%) did not receive the intervention. Of the 100 initial control villages, 8% received the intervention despite their control status.

Linnemayr and Alderman used the IV, DiD and PSM techniques to account for potential biases due to the imperfect execution of random assignment. As the instrument, they used “planned treatment status”. Such a variable was exogenous by construction, and was also a strong predictor of actual treatment. In the first stage of analysis, planned treatment status and other village-level variables were used as instruments for actual receipt of the intervention. In addition, they included initial village-level characteristics such as distance to the next village, prevalence of female education, or the presence of a market that may have influenced the NGOs' placement decision as well as five interactions of village characteristics and the planned treatment. The result of the IV estimation, however, showed the insignificant effect of program actual implementation. The estimator based on the combination of PSM[17] and DiD findings showed a statistically significant result at the 5% level and about a 0.27 standard deviation increase in weight-for-age score.

17 They constrict sample to observations with a propensity score lying in the interval [0.05;0.95].

While the results of the IV analysis showed no significant effect on the intervention, the result from the PSM and DiD analyses appeared to be significant. It might be the case that since the baseline used in the DiD analysis is in the absence of treatment, both groups would experience a similar trend in the outcome variable; in other words, the time trend of the treatment group and comparison group were the same, and there were unobserved time-varying characteristics affecting the results. Also the PSM technique used estimates based on the observed characteristics, and therefore, remaining unobservable differences might have still leaded to a biased estimation of treatment effects. However, The IV estimate may have absorbed those remaining unobserved characteristics. Unlike the follow-up evaluations for the health targeted intervention, the following studies emploed a shock measure such as exposure to war and natural disasters as their instruments.

Alderman, Hoddinott, and Kinsey (Alderman, Hoddinott and Kinsey, 2006) estimated the effect of preschool height on years of completed schooling and delayed enrollment in Zimbabwe. The data consisted of a sample of 665 Zimbabwean young adults in 2000, which was a follow-up to two earlier surveys of children carried out in 1983-1984 and 1987. The authors linked exposure to transitory shocks experienced by children before age three to their pre-school nutritional status, as measured by height-for-age. Education outcomes were as measured as a number of grades attained and the timing of starting

school. Using the combination of maternal FE and IV estimation to showed that exposure to civil war and drought shocks in early childhood in Zimbabwe, the results suggested that reduced height-for-age, which subsequently had a negative impact on years of schooling and adult height. In addition, exposure to civil war and drought were associated with decreases in child height-for-age by 0.049 (t=3.17). The result of maternal fixed effect and IV estimates showed that improvements in height-for-age in children under five leads to increased height as a young adult, and an increased number of grades of schooling completed by 0.678 (t=2.13) as well as being causally associated with starting school at a younger age (0.400 (t=1.65)). Over-identification tests confirmed that these instruments did not violate the exclusion restriction.

The previously reviewed studies have separated the impact of child's nutrition status on schooling from other factors that determine schooling and nutrition status, however the latent problem these studies had was that they did not take into account the importance of early childhood health status. The measures for health used in their studies were after the critical development periods, roughly from age zero to two. Alderman, Hoogeveen, and Rossi (Alderman, Hoogeveen and Rossi, 2009) examined the effect of early child's malnutrition on delayed school entry and the total years of schooling attainment using a panel data set that combine the 2004 Kagera Health and Development Survey

(KHDS) with all individuals who were household members of the KHDS 1991-1994.

They measured health status as height-for-age for children under ten years old in 1991-1994 thus including the critical periods of child's development. However, in order to account for the non-linearity of the impact of nutrition on education output, they used a quadratic form of height which is measured as a percentage of the median of the reference population[18]. This made interpretation of their results very convoluted. Education outcomes were measured as delayed enrollment defined as a child had not entered school at the age of 7 and years of schooling at the second round.

They used community fixed effect to rule out the possibility of bias from unobserved community level effects including infrastructure. They used crop loss happened in 5-6 years prior to the first round interview for controlling for household level characteristics and whether shock experienced 5-6 years prior to the interview for the community level characteristics as the instruments for height-for-age and these variables were interacted with age 0-3 and 4-5 and gender (see the same approach taken by Alderman Hoddinott and Kinsey 2006).

Over-identification tests confirmed that these instruments did not violate the exclusion restriction. With the same limitation applied to Alderman et al (Alderman et al.,

18 The authors take 85% of the median as reference height because stunted children exhibit, on average, a value of height which is equal to 85% of the reference median.

2001), not controlling for the variation within household, their results suggested that one unit increase in height-for-age as a percentage of the median of the reference population percentage decrease 0.591(z=3.42) delayed enrollment and one unit increase in height-for-age as a percentage of the median of the reference population increased the 0.575 years of schooling (z=2.18). However, these results did not give us any intuitive understanding of what these numbers meant. Making attained results more relevant in terms of change in height-for-age, the authors run simulation for the results and showed that if the height of a boy was increased from 80% to 95%[19] of the median, this would increase schooling by 0.93 years, while the increase was 0.85 years for boys whose height was equal to 85% of the median. In fact, although not shown in the paper, these results were similar to the results for OLS which were the years of schooling for boys at 85% median is 6.9 and 7.05 for 90%[20].

In recent years, several new longitudinal studies have investigated the effect of childhood malnutrition on education outcomes such as school delayed school entrance, final grade attainment and test score (Victora et al., 2008; Jones et al., 2008; Glewwe and Jacoby 1995; Glewwe, Jacoby, and King 2001; Alderman et al., 2001). Few results from

20 The reason for this is that they include rich covariates in their OLS regression. The covariates in the analysis include age of child, gender, household expenditure, parent having secondary schooling, maximum grade of parents, heights of parents, school quality as indicated by both teacher pupil ratio and boards per class, urban or rural, and access to electricity.

quasi-experiments which employed a shock measure as exposure show that increased health status of a child had a causal effect on education outcome such as delayed enrollment and final grade attainment (Alderman, Hoddinott, and Kinsey 2006; Alderman, Hoogeveen, and Rossi 2009, Linnemayr and Alderman 2011). Earlier studies using cross-sectional data had already shown that an increasing level of stunting shows negative association to education outcomes (Grantham-McGregor and others 2007, Glewwe and Jacoby 1995).

In order to address endogeneity, the most of these studies used FE approach to control the unobserved household characteristics which do not change over time and do not interact with variables that do change over time. The IV method was used to account for the bias due to correlation of the observed variables with that error term other than household characteristics. A combination of two approaches seems valid only if household unobserved characteristics such as parental teats do not vary within same household.

In a quasi-experimental design, in addition to the IVFE approach, DiD and PSM were used in the reviewed literature. Since previous literature used an exogenous shock measure, the results were more likely to estimate a true effect size. However, for the studies with the initial program intervention, the common problem with randomized trials was that there was a departure from the initial design which ultimately biases the results. Especially

an adherence to originally assigned treatment status became increasingly difficult in reality. The results from these empirical studies controlling for endogeneity are consistent in the sense that improved child health, as captured by height, appeared to have a large positive causal impact on education outcomes. However, once controlling for range of covariates, the results attained from “naïve” estimates show the very similar results as improved estimates accounting for endogeneity (see Alderman, Hoogeveen, and Rossi 2009).

2.4. Conclusion

The association between the timing of the Kasai procedure and survival of native liver has been suggested in many studies; however, these studies did not fully account for unmeasured confounding.

Three studies (Lien et al., 2011; Chardot et al.1999,; Chardot et al., 2013) conducted multivariate analyses, and in their studies, the early timing of Kasai procedure was associated with better prognosis.

Yet, none of the previous literature focused on the threshold of the timing or the causal effect of the Kasai procedure. The possible categorization for the timing of Kasai procedure will be less than 30 days and added up by 15 days since the earlier intervention seemed to larger impact on the later outcomes in France study with the largest sample size

in the existing study (Chardot et al 1999, 2013). T

The review on the literature on investigating the causal effect of health on education revealed that many of previous studies in economics and education have used the quasi-experimental techniques such as IVFE approach, DiD and PSM to address the potential biases in observational studies. The most of these studies used a FE approach to control the unobserved household characteristics which do not change over time and do not interact with variables that do change over time. The IV method was used to account for the bias due to correlation of the observed variables with that error term other than household characteristics. A combination of two approaches seemed valid only if household unobserved characteristics such as parental teats did not vary within same household.

By adapting the methodologies developed and used in economics and education, I will estimate the causal effect of the timing of the Kasai procedure on patient outcomes. In order to apply the methodologies used and developed in economics and education into the causal analysis of the timing of the Kasai procedure in BA study, in the next chapter, I will discuss the conceptual framework of causal inference.

Chapter III

CONCEPTUAL FRAMEWORK

3.1. Introduction

The goal of this dissertation is to provide comprehensive evidence on the effect of the timing of the Kasai procedure on patient outcomes. Since none of previous literature has focused on the causal effect of the timing of the Kasai procedure on patient outcomes or trying to control for rich sets of covariates[21], this study will shed lights on many important features of BA that have not been fully addressed before.

My main interest is: 'what is the expected outcome to patients by change in the timing of the Kasai procedure relative to the counterfactual situation?' The focal point of the question is to identify the factors that contribute to variation in the timing of Kasai procedures in infants with BA. While the randomized control trial is considered the gold standard for the evaluation of the effect of a particular treatment on patient outcomes, in many settings (including the one presented with my research question), it is impractical and unfeasible to measure outcomes in a complete "counterfactual" scenario, as a strict treatment assignment is difficult to pursue. Since the negative relationship between the delayed the Kasai procedure and prognosis has been previously suggested (Lien et al.,

21 See Chapter 2

2011; Chardot et al.,1999; Chardot et al., 2013), the designing randomized controlled trial is not the most desirable way to evaluate the causal effect of the early operation of Kasai.

Many previous studies in economics and education have developed quasi-experimental techniques to address the potential biases in observational studies. Since the strength of this study is that the data with the largest number of subjects in the world, 2743 patients allowed overcoming the limitations with the small sample size, which many of previous studies encountered.

In order to understand the potential biases and approaches to overcome these biases, it is important to understand the foundation of causal inference and threats to causal inference, particularly the issue of selection bias and endogeneity in observational studies. After describing the potential biases, I will suggest four statistical approaches that could potentially mitigate the effects from such biases.

3.2. Foundation of Causal Inference

Angrist et al. (Angrist, Imbens and Rubin, 1996) explained that causal inference in statistics evolved from work by Fisher (Fisher, 1919) on agricultural experiments. Fisher's work on the basic formulation of causal inference was then extended by Rubin (Rubin, 1974) providing greater emphasis on more complicated scenarios such as those which were

often the focus in observational studies without randomization. This approach is now widely used, and is typically referred to as either the potential outcome approach, or the Rubin Causal Model (Holland, 1986). In the basic form of this model, there are two potential outcomes (Y^T, Y^C) for each individual, where Y^T indicates an outcome of an individual who is assigned to the treatment (such as an early the Kasai procedure) and Y^C indicates the outcome of a similar person without a given treatment. Further we define a binary assignment indicator D, indicating whether an individual actually received the given treatment (D = 1) or not (D = 0). The treatment effect for each individual is then defined as the difference between potential outcomes:

$\Delta = Y^T - Y^C$--- (1)

As described in equation (1), the causal effect of the treatment on a targeted subject is the difference between the outcome under treatment and without treatment (the counterfactual scenario). The fundamental problem of causal inference persists in the fact that the outcome cannot be observed for both situations for the same individual. Instead, the average of the entire sample is taken, and the quantity is noted as:

$E(Y^T - Y^C)$ --- (2) , and called Average Treatment Effect (ATE)

In the experimental settings, individuals are assigned randomly to a treatment group or control group. Successful randomization allows us to see the pure effect of the treatment on those who are assigned to the treatment can be estimated through the

evaluation parameter, ‘average treatment effect on the treated’. It ensures statistical equality of observed and unobserved characteristics, and is therefore independent from potential outcomes and assignment to the treatment. The assumption that the treatment effect Δ in equation (1) for each person must be independent of the treatment of other individuals and this is called the stable unit treatment value assumption (SUTVA). The parameter for estimating the population average of gains from treatment is the so-called ‘average treatment effect on the treated’:

$$E(\Delta | D = 1) = E(Y^T | D = 1) - E(Y^C | D = 1) \text{ --- (3)}$$

It could only answer the question of ‘what is the expected or mean outcome gain to individuals who received treatment to the hypothetical situation had they not received it?’ The second term in (3) is the hypothetical outcome without treatment for those subjects who received treatment. If the condition

$$E(Y^C | D = 1) = E(Y^C | D = 0) \text{ --- (4)}$$

holds, then the non-assigned patients can be used as a control group. The above-mentioned condition is the ideal situation where there is a random allocation in which to take part. However, when randomization is not possible, there are threats to causal inference, and therefore the ability to assess the real effect of the intervention is in question. . The following subsection will explain the potential threats to causal inference.

3.3. Threat to Causal Inference

In a real-world setting, there are many situations where it could be harmful to assign or not assign a given medical intervention, or it could be the case that the considerable cost of designing and implementing a randomized controlled trial would make such a study unfeasible.

3.3.1. Experimental Study

Selection bias occurs when study subjects self-select into either the treatment or the control group based on certain characteristics (or physicians select patients due to particular characteristics). Some patients take the treatment because it is believed their response to the treatment is quite large, whereas those who do not take the treatment have either minimal response or even a negative response to the treatment. If there is heterogeneity in the response to the treatment, the calculation $E(Y^T| D = 1) - E(Y^C|D = 0)$ is insufficient to estimate the effect of the treatment. The randomization design does not remove such selection bias, but balances it between the treatment and control groups, effectively cancelling out its impact on the observed treatment effect. Randomization ensures that, on average, both groups are statistically equivalent in both observed and unobserved characteristics, and therefore, the potential outcomes are independent of the

assignment to the program. However, as stated above, randomized control studies are not always feasible in social science settings, and therefore the causal effect of the treatment cannot be calculated as described in equation (1). Instead, the observed outcome for each individual is given by:

$Y = D * Y^T + (1 – D) * Y^C$ --- (5)

Equation (5) shows that the two potential outcomes, Y^T and Y^C, cannot be observed for the same individual simultaneously. The unobservable component in (5) is called the counterfactual outcome, so that for individuals who took the treatment (D = 1), Y^C is the counterfactual outcome, and for those who did not (D = 0) it is Y^T. In this setting, the problem of evaluating the individual treatment effect can be interpreted as a missing data problem because for any given individual the counterfactual outcome cannot be estimated. However, there are limitations to non-randomized study designs, as there are situations where it appears one cannot actually measure both Y^C and Y^T at the same time.

3.3.2. Observational Study

The latent problem of selection bias persists more in non-experimental studies since these studies do not allow such random assignment to occur. In addition to the selection bias, it is more likely that the observed independent variable and explanatory

variable are correlated with unobserved or confounding variables influencing both the early Kasai procedure and outcomes in observational studies. This problem is referred to as endogeneity, and this can become a problem in a typical empirical model. The simple linear approximation of the effect of early Kasai procedure on patient outcomes can be expressed as:

$p(SNL)_i = \beta_0 + \beta_1 K_i + \beta_2 X_i + \alpha + \mu_I$ --- (6)

where SNL_i is a survival with native liver, K_i is the early Kasai procedure , X_i is a set of a vector of other inputs that influence survival of native liver, α is the child's innate healthiness that would account for the severity of the disease, and μ_i is a normally distributed error term. The term β_1 is interpreted as the effect of "early Kasai procedure" on outcomes of patients. In observational studies, the problem arises when there are unobserved factors that affect both the survival with native liver and the timing of Kasai. It is the nature of rare disease that there are many unknown factors that might affect both the predictor and the outcome. In addition to such uncertainty, there would be "potentially observable" characteristics but not taken as variables therefore not in the analysis. Such important determinants may include any characteristics relating to the sociodemographic and socioeconomic status of patients, as well as environmental (geographic and seasonal), clinical, hospital and other potential factors such as the characteristics of each local government providing or influencing health

services for childcare.

Much of the available literature has focused on the clinical characteristics of patients in relation to the timing of the Kasai procedure and outcome. For example, types of obstruction (Nio et al., 2006) and existence of malformation (Shneider et al., 2006; Dillon et al., 1994; Davenport, 2006; Schwarz et al., 2013) were proven to be predictors for later improved prognosis. For the hospital characteristic, it appeared that a center that performed at least five the Kasai procedures per year had a better success rate, as measured by 5- or 10-year long-term survival with the native liver (Davenport et al., 2011; Chardot et al., 1999). However, the results noted above come from studies that did not fully control for any potential confounders. While certain potential confounding factors may not have an obvious influence on short-term outcomes, they may prove to be pertinent later.

Other characteristics could include the local government's involvement or influence on the provision of child health services in general, which could vary substantially, particularly in Japan. Some local governments provide home visiting by nurses to check the general healthiness of infants and mothers at one month after a delivery. Since the early detection of BA is largely depending on catching the early symptoms such as persistent jaundice, home-visiting by medical professional might influence the early detection leading to the earlier the Kasai procedure. Also as explored in the previous studies, geographic and

seasonal variation might affect the outcomes as well.

The innate healthiness of a child, α in the basic model, is difficult to measure. Although often birth weight is used as a proxy for the innate healthiness in more general settings when investigating the effect of health on educational outcomes, for the case of BA, which is more disease-specific, the use of birth weight as a proxy for the innate healthiness might be misleading because patients are usually born at full term. Measuring motivations among medical staff and parental characteristics on the management of the whole process are very difficult in reality. Also parents' focus on a child and ability to manage the pre-and-post operation are difficult to measure as well. There are both observed and unobserved processes that lead to determining the timing of the Kasai procedure, and these factors are correlated with survival with native liver, either directly or indirectly.

3.4. Empirical Approach to the Problems of Selection Bias and Endogeneity

After briefly explaining the traditional logit regression approach, both weakness and strengths of four approach will be disused. These approaches will be DiD, PSM, FE and IV approach in addition to multivariate logit regression as the base model.

3.4.1. Logit Regression: The Basic Model

The following shows the basic empirical model to estimate the effect of early Kasai intervention on the outcome:

$\log(p/(1-p))SNL_{kt} = \alpha_0 + \beta_1 K_{kt} + \beta_2 X_{kt} + \alpha_{kt} + \varepsilon_{kt}$ --- (7)

where SNL_{kt} is a probability of survival of native liver at t (1-, 5-, 10-, 15-, and 20-year) with k individual. K_{kt} is the timing of the Kasai procedure. X_{kt} is a vector of other inputs that influence survival of native liver. According to the previous studies, anatomic pattern (Nio et al., 2006), existence of malformation (Shneider et al., 2006; Dillon et al., 1994; Davenport, 2006; Schwarz et al., 2013), and expertise of the surgeon and care center had strong associations with the outcome (Dillon et al., 1994; Davenport, 2006). These characteristics should be included in X_{kt}. The error term ε_{kt} captures unmeasured variables.

Although in this study, rich set of covariates including anatomic pattern, existence of malformation, and expertise of surgeon and care center, there still would be a latent problem of omitted variables. α_{kt} indicates the child innate “healthiness” that would affect the progress and severity of the disease. While various sociodemographic and clinical characteristics available in this study and these might be partially able to represent the innate healthiness, realistically it is very difficult to control in the model. There is potential bias in this model due to possible correlation between K_{kt} and ε_{kt} resulting from any

correlations between sociodemographic, socioeconomic, clinical, and hospital characteristics in addition to child's innate healthiness. However, this potential bias can be corrected for by using IV and FE techniques.

In multivariate logit regression, it is assumed there is no left-out variable or confounding variable that may bias the estimates associated with the effect of the timing of the intervention, in this case, the effect of a Kasai on patient outcomes. If this condition does not hold, then the estimates of K_{kt} using the logit regression will be biased. But in reality, it is always difficult to include all possible covariates in a model. In order to mitigate such bias, DiD, PSM, FE, and IV have been used in the previous studies performed in other settings to see the effect of health on education.

3.4.2. Difference-in-Difference

Linnemayr and Alderman (2011) use DiD and PSM in their analysis. The DiD approach can generate a reliable estimate in scenarios where there is a natural experiment such as a sudden policy change. The DiD estimation method can be interpreted as an extension to the classical before-after estimator. In order to compare the differences between before and after, we need have a "shock" measure, which acts like an intervention in the experimental design. Using "exposure to shock" as a health shock that acts like an

intervention in a natural experiment, DiD estimates the effect of such “shock”.

In order to apply this technique, it is necessary to have data on pre and post exposure covering both those who are exposed and those who are not, as well as one or more additional follow-up surveys after the exposure. The following table (Table 8) represents the steps necessary to estimate a DiD estimator. There are four data points, namely pre-exposure (t0) data points ($Y_{t=0}$ and $C_{t=0}$), and post-exposure data points ($Y_{t=1}$ and $C_{t=1}$). The difference of differences between pre- and post-data points is DiD estimator $((Y_{t=1} - Y_{t=0}) – (C_{t=1} – C_{t=0}))$.

Table 8: Calculation of DiD

	Exposed group	Control Group	Difference across groups
T1	$Y_{t=1}$	$C_{t=1}$	$Y_{t=1} - C_{t=1}$
T0	$Y_{t=0}$	$C_{t=0}$	$Y_{t=0} – C_{t=0}$
Difference Across time	D1: $Y_{t=1} – Y_{t=0}$	D2: $C_{t=1} – C_{t=0}$	DiD: $(Y_{t=1} – Y_{t=0}) – (C_{t=1} – C_{t=0})$

The assumption needed for DiD estimator is that the time trend of the treatment group and comparison group is the same. In order to assure that the comparison group is similar to the treatment group before doing the double difference, often the PSM method is used.

Recall Eq (7) which is rewritten as:

$$\log(p/(1-p))SNL_{kt} = \alpha_0 + \beta_1 Z_k + \beta_2 T_k + \beta_3 Z_k T_k + \beta_4 X_{kt} + \varepsilon_{kt} \quad ---(8)$$

where Z_k denotes exposure group (living in area affected by shock or not) and T_i denotes time period (pre- and post-redistricting). Z_kT_k is an interaction, which measures precisely the differences in means across two levels of a dummy variable for the only observations that actually get exposed.

In my analysis, the implementation of stool color cards was intended to be included in this analysis because it stands as a shock measure. The national color card screening was launched in 2012 and prior to that there were 10 regions in Japan where implemented the introduction of a stool color card. The necessary data for this technique was data on pre and post exposure, which must cover both those who were exposed and those who were not, and one or more additional follow-up surveys after the exposure should be in place. Recall Eq (8), where Zk denotes the exposure group (living in an area where a pilot screening was implemented) and Ti denotes the time period (pre- and post-pilot studies; 1=before year X 0=after year X). ZkTk is an interaction, which measures precisely the differences in in means across two levels of another variable that is a dummy variable for the only observations that actually get exposed (born before/after or patients who lived in 12 regions in Japan after year X).

3.4.3. Propensity Score Matching

Linnemayr and Alderman (2011) used PSM to assure that the comparison group is

similar to the treatment group before applying the double difference approach, (as explained in the previous sub-section).

Recall Eq (1), there are two potential outcomes (Y^T, Y^C) for each individual, where Y^T indicates an outcome of an individual who is assigned to the treatment, and Y^C indicates the outcome of a similar subject without a given intervention. Further, a binary assignment is defined as D, indicating whether an individual actually participated in an intervention (D = 1) or not (D = 0). Rosenbaum and Rubin (1983) suggested selecting a propensity score at random from the sample, then matching two individuals with this propensity score (one who chose treatment and one who did not), and constructing their outcome difference as:

$Y^T - Y^C - |p(X)$---(9)

If the condition $E(Y^C|D = 1) = E(Y^C|D = 0)$ --- (4) folds the expected difference is $E(Y^T - Y^C|p(X))$ Average Treatment Effect (ATE) conditional on $p(X)$.

This assumption is called conditional independence assumption or ignorability. In other words, $E(Y^T - Y^C|p(X))$ is conditional on the observable set of confounders, and there is no selection bias based on unobserved heterogeneity. If this strong assumption holds, then the two potential outcomes are independent of the assignment to treatment, conditional on X. Although, this is a strong assumption and has to be justified by the data quality at hand. The propensity score can be calculated by using a standard logit or probit model, where the outcome is the treatment indicator, and the predictors are all observable variables. It can

occur that controlling for selection on observables may not be sufficient since remaining unobservable differences might still lead to a biased estimation of the treatment effect. The ideal data required for this method would include a large number of subjects and with more variables to facilitate good matching.

3.4.4. Fixed Effect

The fixed effect method is also used frequently in combination with the use of IV in the reviewed literatures (Alderman and others 2006; Maluccio, Hoddinott, and Behrman 2006; Behrman and others 2006; Glewwer and Jacoby 1995, 2001). In this study, the question is 'what is the effect of timing of the Kasai procedure on patient outcomes?' Intuitively, to estimate this effect, one has to account for the fact that hospitals might differ in characteristics. More specifically, the expertise of surgeon among other characteristics of each hospital would certainly result in positive or negative correlations between the timing of the Kasai procedure and outcomes that are not directly causal.

Recall the equation (7), I can estimate the parameter τ from the regression:

$$\log(p/(1-p))SNL_{kt} = \alpha_0 + \beta_1 K_{kt} + \beta_2 X_{kt} + \tau Z_{kt} + \gamma_{k\,t}\,\varepsilon_{kt} \quad \text{--- (14)}$$

where $E(\varepsilon_i | X_1, \ldots, X_K, \gamma, Z) = 0$. Since γ indicates unmeasured covariets, it is impossible to include in the equation. The possible solution for this is that suppose we observe $X_1, \ldots, X_K,$

Z for two (or more) “paired” observations with the same value of γ. To be concrete, suppose there are *2n* observations. Unlike the case above, there is no need to estimate a separate γ parameter for each unit, just for every pair of units. When the number of paired observations exceeds the number of covariates, it is possible to estimate all the model parameters. The hospital component (hospital fixed effect) and are of residence (geographical fixed effect) are possible candidates.

There are limitations associated with the FE method, as seen in Glewwe and Jacoby, that FE cannot counted for variations in innate child healthiness or random shocks to timing of the Kasai procedure among patients within the same hospital. FE model controls variations within hospital. The previous literature has suggested using both FE and the IV because IV complements FE by controlling for variations across hospital. This combined approach has been applied by many researchers (Alderman and others 2006; Maluccio, Hoddinott, and Behrman 2006; Behrman and others 2006; Glewwer and Jacoby 1995, 2001).

3.4.5. Instrumental Variable

The IV approach relative to logit regression, DiD, PSM and FE, which do not

control for unmeasured confounding is that IV approach could potentially provide a way to control for unmeasured confounding. Many of reviewed studies examining the effect of health on education used the IV approach (Alderman and others 2006; Maluccio, Hoddinott, and Behrman 2006; Behrman and others 2006; Glewwer and Jacoby 1995, 2001).

Idealistically, since the timing of Kasai procedure ($K_{kt)}$ might be endogenous, we can get an IV estimate of $\beta_1$as:

$$b_1^{IV} = \frac{Cov(y, z)}{Cov(K, z)} = \frac{Cov(\alpha_0 + \beta_1 K + \varepsilon, z)}{Cov(K, z)} = \frac{Cov(K, z)\,\beta_1 + Cov(\varepsilon, z)}{Cov(K, z)} \text{ --- (10)}$$

If the assumptions are correct about the instrument z, then Cov(ε, z) = 0, and therefore:

$$b_1^{IV} = \frac{Cov(K, z)\,\beta_1 + Cov(\varepsilon, z)}{Cov(K, z)} = \beta_1 \text{ --- (11)}$$

So the IV estimator will be an unbiased estimator β1. Using this IV, the first stage in the two-stage least square (2SLS) is given by:

1^{st} Stage: $K_{kt} = \gamma_0 + \gamma_1 z_{kt} + \upsilon_{kt}$ ---(12)

where K_{kt} is a regressor: timing of the Kasai procedure, and z is an IV. In Eq (12), it fits an OLS regression model to the hypothesized relationship between K and z as where the first-stage parameters γ_0 andγ_1 represent the requisite intercept and slope, and υ_{kt} is the first-stage residual that contains that part of question K that remains unpredicted at the first

stage. After obtaining the required predicted values from the first-stage fit, the second stage of the 2SLS becomes:

2^{nd} Stage: $pSNL_{kt} = \alpha_0 + \beta_1 \widehat{K}_{kt} + \beta_2 X_{kt} + \varepsilon_{kt}$ ---(13)

There are certain assumptions that allow for validation of the instrument that randomizes observable characteristics relevant to the outcome. This suggests that they are expected to also randomize unobservable confounders. Angrist et al. (Angrist, Imbens and Rubin, 1996) considered an IV to be a variable which satisfies five assumptions: SUTVA, the IV is positively correlated with treatment assumption, the IV is independent of unmeasured confounders assumption, the exclusion restriction assumption, and the monotonicity assumption. The SUTUVA assumption states that there are no different versions of the treatment that have different effects and there is no interference between units. Also the last monotonicity assumption states that a hypothetical change in the instruments either has no impact on a unit's treatment status, or changes its treatment status in the same direction as it does for all other units for which it has an impact. Monotonicity is also plausible in this study that the encouragement (Z =1) provides a clear incentive to take the treatment. Among five assumptions, the core assumptions of IV are depicted in Figure 7 which I will also be carefully examined in Chapter 4 as well.

Firstly, the correlation between z_{kt} and K must be non-zero in the equation (13). In

other words, the instrument must be related to the potentially endogenous question predictor. In Figure 7, it is expressed as the arrow from Z to D. The second important condition that must be satisfied for successful IV is that the IV is not associated with variation in unobserved variable. z_{kt} is uncorrelated with ε_{kt} in the equation (13) and it is represented by the crossed-out relationship between Z to U in Figure 7. The condition is also appealing conceptually since if the instrument were correlated with the residuals in the original equation, it would suffer from the same problem as the original OLS estimate. However, this condition is difficult to be confirmed because we cannot observe ε_{kt}. The third condition is the instrument (z) only affects the outcome through its effect on the treatment. This means that a good IV should to be highly correlated with a regressor but uncorrelated with ε_k in the equation (13). It is represented by the crossed-out line from Z to Y and Z to D to Y in Figure 7.

In summary, even though the fixed effect and IV approaches were used in the reviewed literature, there are other approaches including DiD and PSM approaches that might be possible to apply in the study. The reason why other approaches are not so popular among researchers could be the reasonability of holding the required assumptions of each approach. The assumption needed for DiD estimator is that the time trend of the treatment group and comparison group is the same. PSM can only be based on the

observed characteristics of patients. At the same time, the reasons why the fixed effect accompanied by IV methods are used more often are that the hospital fixed model could control variations within hospital and the IV controls for variations across these differences. IV is the only approach for mitigating the effect of bias caused by endogeneity.

3.5. Conclusion

The ultimate goal of this study is to estimate the causal effect of the Kasai procedure on patient outcomes. The fundamental problem of causal inference persists in the fact that the outcome cannot be observed for the counter factual scenario for the same individual, and therefore, one must observe the average treatment effect of the entire sample. Experimental settings with successful randomization allow us to see the causal effect of the treatment on those who are assigned to the treatment through the evaluation parameter, average treatment effect on the treated. But since this study used observational data, the random allocation to the treatment cannot occur, and therefore, potential threats to causal inference such as selection bias and endogeneity exist to accurately estimate the treatment effect.

Multivariate logit regression with rich sets of covariates, DiD, PSM, FE, and IV were reviewed for potentially mitigating bases. The problem with observational data is that there

are an infinite number of unobserved variables that could make the observed relationship endogenous. There are certain assumptions underlining in each approach. The core assumption associated with the DiD approach is that the time trend of the treatment group and comparison group is the same. For the PSM, there is a conditional independence assumption that that treatment status is random, conditional on some set of observed covariates. The FE approach assumes that the unobserved hospital characteristics, which do not change over time and do not interact with variables that do change over time. However, these methods only control for measured confounders and do not control for unmeasured characteristics. The IV method is therefore useful in conjunction with these other methods, because it has the potential to control for unmeasured characteristics. The core assumptions for the IV method are that the instrument must be related to the potentially endogenous question predictor, the instrument is independent of unmeasured confounders, and it affects outcomes only through its effect on treatment.

Chapter Ⅳ

METHODOLOGY

4.1. Introduction

None of previous literature in BA has addressed the problem of endogeneity when aiming to assess the effect of the timing of the Kasai procedure on patient outcomes. Moreover, only three studies based on the population-based datasets tried to control for the confounding variables (Lien et al., 2011; Chardot et al.,1999; Chardot et al., 2013). Their results suggested that in addition to the timing of the Kasai procedure, the anatomic pattern of BA, existence of associated anomalies, prophylactic antibiotics, and hospitals with high caseload independently associated with patient outcomes.

Drawing from research literature in the fields of economics and education, four approaches were identified in order to correct for confounding in quasi-experimental research designs. The four approaches identified for use in this study were the DiD, PSM, FE, and IV methods.

4.2. Data Description and Sample Selection

This study used the data drawn from the Japan Biliary Atresia Registry (JBAR). In

1989, the Japan Biliary Atresia Society launched a nationwide registry to investigate all aspects of BA. A total of 2,743 patients (1,743 girls, and 1,000 boys) have been registered since then. The JBAR dataset includeed BA patients who were diagnosed as BA from 1989 to 2012. The JBAR was based on the observational data gained through initial and follow-up questionnaires.

The variables in JBAR included the sociodemographic background of patients, clinical and hospital characteristics. While the JBAR contained a lot of variables that have not been investigated previously, this study utilized a sub-set of these because some of them were still under the data cleaning process.

The JBAR had a high rate of a national BA population coverage. While it is impossible to know the exact number of all BA patients due to the mortality rate associated with the disease, an estimate of the number of BA patients living in Japan is provided below in Table 8. According to the calculation, for an incidence of 1 in 9640 (Nio et al., 2003), the estimated total number of BA patients in Japan from 1989 to 2012 would be 2864. Therefore, the total number of patients captured in the JBAR from 1989 to 2012 was 2743, suggesting that the population coverage of JBAR was high (96%). This indicated that selection bias due to not representing the target population was not likely a major source of bias in this study. The use of a national population registry allowed for minimization of selection bias, and

could facilitate greater generalizability of the study findings to the target population.

The inclusion criteria for this study stipulated that the subjects were born between 1989 and 2012 and underwent a Kasai procedure. Also I only included patients in whom I could observe the primary predictor and outcome. However, the outcome was not observed or unknown for 122 patients identified in the registry, and therefore these patients were excluded from the study. There were also additional 16 subjects who did not receive a Kasai procedure, and were therefore also excluded.

The primary predictor variable used in this study was the age at Kasai procedure and the outcome was 1-year survival of native liver, and data for both and were present in a total of 2605 subjects. The timing of Kasai procedure was broken into categories starting the less than 30 days, and incremented by 15 day periods. The number of subjects receiving the Kasai procedure less than 30 days was 158, 525 for less than 45 days, 1147 for less than 60 days, 1,780 for less than 75 days, 2175 for less than 90 days, 2,356 for less than 105 days and 2,449 for less than 120 days.

Other variables included sociodemographics including gender, birth order, birth weight, gestational age, parental age, clinical characteristics including type of obstruction, associated anomalies, and type of choleretic agents, and hospital characteristics including a hospital with a high caseload. Types of obstruction were defined according to the Japanese

Society of Pediatric Surgeons classification[22]. Associated anomalies were defined as any indication of situs inversus, asplenia, polysplenia, accessory spleen, and preduodenal portal vein[23]. Type of choleretic agents include dehydrocholic, ursodeoxycholic acid, secretin, steroid, PGE1, taurine, PG-F2α, and PG-E2. Hospitals with high case load was defined as hospitals which had the mean caseload (59) +1SD (31) cases.

4.3. Estimation Strategy

The assumptions in the empirical analyses can be based on two perspectives: the first that there were no unmeasured characteristics in the model, and the second was that instead, that there were some unmeasured characteristics left out from the model.

Existing literature (Lien et al., 2011; Chardot et al., 2013) supported the idea that, a younger age at the time of the Kasai procedure consistently has been found to be associated with better outcomes. These empirical studies supported that early diagnosis had an impact on later outcomes as a pathway through the early intervention of the Kasai procedure. Recall the following shows the basic empirical model to estimate the effect of the early Kasai intervention on the outcome.

$$\log(p/(1-p))SNL_{kt} = \alpha_0 + \beta_1 K_{kt} + \beta_2 X_{kt} + \alpha_{kt} + \varepsilon_{kt} \text{ --- (7)}$$

22 See Figure 2 for the description of types of obstruction
23 See Table 1 for the description of associated anomalies and choleretic agents

where SNL_{kt} is a probability of survival of native liver at t (5-, 10-, 15-, and 20-year) with k individual. Although including the long-term outcomes in the analysis was my ultimate interest, the data in hand only allowed examining the effect on the short-term outcome, 1-year native liver survival without jaundice. A native liver survival without jaundice has been considered as good prognosis, and its association with better long-term prognosis has been suggested in prior research (Lien et al., 2011). The independent variable in this analysis was the timing of Kasai procedure. K_{kt} is the timing of the Kasai procedure. X_{kt} is a vector of other inputs that influence survival of native liver. α_{kt} indicates the child innate healthiness that would affect the progress and severity of the disease. ε_{kt} is an error term that captures unmeasured variables.

According to the previous studies, anatomic pattern (Nio et al., 2006), existence of malformation (Shneider et al., 2006; Dillon et al., 1994; Davenport, 2006; Schwarz et al., 2013), and expertise of the surgeon and care center had strong associations with the outcome (Davenport et al., 2011; Chardot et al., 1999). These characteristics should be reflected in X_{kt},

4.3.1. Bivariate Analysis

I conducted a bivariate analysis to investigate the direct relationship between the

outcome and the timing of Kasai procedures and other covariates. This process was important for two reasons: First of all, the most of previous literature only looked at bivariate relationship between the timing of the Kasai procedure and outcome; therefore by conducting the bivariate analysis, the result would be comparable to the previous studies. The results were also compared to the results from the multivariate analysis and examined the difference among them.

4.3.2. Multivariate Analysis

In the multivariate analysis including logit/probit regressions, it was assumed there was no left-out or confounding variable that may generate the biased estimates associated with the effect of the timing of Kasai procedure on patient outcomes.

The covariates used in the multivariate analysis included sociodemographic characteristics (e.g. gender, birth order, birth weight, gestational age, parental age), clinical characteristics (e.g. type of obstruction, associated anomalies, and type of choleretic agents), and hospital characteristics (e.g. hospital caseload).

Since the dependent variable in the analysis (1-year native liver survival without jaundice) was dichotomous, the relationship between the dependent variable and independent variables was non-linear because the probabilities were bound between 0 and 1. Even though binary outcomes are a result of a latent variable which is continuous and

normal, the results from probit regression can be difficult to interpret. Therefore, in this study, the results of the two regression models, logit and probit regressions were compared.

The appropriate selection of potential confounders and inclusion into the model has been widely discussed in the epidemiologic literature (Budtz–Jørgensen et al., 2007; Maldonado and Greenland, 1993). A decision on the best set of variables in the model might be based on the clinical or statistical significance. In this study, first I examined that whether statistical and clinical significance of covariates were previously suggested or not and examined that the variables were statistically significant in the bivariate analysis. The study further examined the inclusion of the variable leads to a 10% change in the main effect in the logit regression. The selection rule was that if the variable was suggested to have a previous indication for its significance, statistical significance in the bivariate analysis, or leaded to a 10% change in the main effect when it was added to the model, then it was included.

A ROC was also used as a measure of model fit. The points on the line indicate the probability of correctly predicting the outcome. It covers all possible thresholds (cut-off points) and the closer the ROC plot is to the upper left corner, the higher the overall prediction (Zweig and Campbell, 1993).

Since confounding does not solely depend on statistical significance or the 10% rule[26], a cut-point for defining confounding has not been fully established. Budtz–Jørgensen et al., 2007) provides a good example in their statistical analysis section describing how they analyzed raw test scores adjusted for previously suggested confounders. Their covariates were included in the full model if the *p*-value was less than 0.20 or if their inclusion resulted in a change of 10% or more in the estimate of the main effect of the outcome. Following the suggested rules, this study compared the predictability of the outcome in the model with full covariates and also the model with selected covariates.

4.3.3. Fixed Effect

The reason for using FE was to address the fact that the characteristics of hospital would affect both the timing of the Kasai procedure and the outcome.

Recall the equation (7), one can estimate the parameter τ from the regression:

$$\log(p/(1-p))SNL_{kt} = \alpha_0 + \beta_1 K_{kt} + \beta_2 X_{kt} + \tau Z_{kt} + \gamma_{k\,t}\,\varepsilon_{kt} \text{ --- (14)}$$

where *Z* for two (or more) “paired” observations with the same hospital.

---

[26] Variables that were not significantly associated with disease severity but that changed the odds ratio for severity by 10% or more when removed from the analysis were also kept in the final model (Maldonado et al., 1993)

The rationale behind FE is that the timing of the Kasai procedure varied within hospitals but any hospital characteristics affecting the timing of the Kasai procedure and outcome were unlikely to vary significantly within each hospital. In this setting, correlation between the timing of Kasai procedure and the outcome should not be detected within a single hospital. My hypothesis, and rationale for using FE regression was that if the variable for hospital characteristics in the multivariate logit regression (defined as the hospital with higher case load [more than 1SD of the mean caseloads of all hospitals]) accounted for the variations in hospital characteristics, then the results between the multivariate logit regression and FE would be the same.

4.3.4. Instrumental Variable

The study performed the multivariate regression with rich covariates and FE to address the scenario whereby no confounding was unobserved. As discussed in detail in section 3.2.5., there is the problem of latent endogeneity when estimating the effect of the Kasai procedure on outcomes. Besides the fact that much is unknown about the cause of BA, there would be potentially observable characteristics that are not included in JBAR data. Such variables could be characteristics of medical institutions including expertise and motivation of staff, parents' characteristics relating to the management of a process of screening and post-operation periods, (e.g. adherence to medication), and hospital and

local government characteristics relating to investments in child health services. In other words, there would be a non-zero correlation between the question predictor and residuals in the population. In any of these cases, estimates of the timing of the Kasai procedure using multivariate logit regression or with FE will be biased. In order to account for such endogeneity bias, one strategy is to use IV estimation.

Two possible instrumental variables were considered for the analysis: 1) whether or not patients received any kind of screening prior to the Kasai procedure and 2) length of stay at last hospital admission prior to having the Kasai procedure. For the first instrumental variable in the dataset, there is a variable describing whether or not patients received any kind of screening prior to being referred to the hospital for detail examinations.

Two different functional forms were used in this analysis. As it is stated in Angrist & Pischke's book, Mostly Harmless Econometrics (Angrist and Pischke, 2009), combining IV with either probit or logit is not an optimal approach because the econometrics are more complicated and prone to errors; therefore, I useed both OLS and probit regressions to perform the analysis and compared if there was any difference between two models.

4.4. Conclusions

In the multivariate analysis including multivariate logit and probit models, it is assumed there are no unobservable characteristics that change over time and can affect patient outcomes in the ways explained that are associated with selection and endogeneity biases in the estimation strategies used in the analyses. Potential problems associated with child's innate healthiness and hospital characteristics are ruled out in the analyses by controlling for birth weight and the indication of hospital with high caseload in the model. The validity of controlling the hospital effect with the covariate indication the hospital with higher caseload is compared with the hospital FE. If there is an endogeneity issue, then the IV should be taken.

The reviewed assumptions are needed in order to argue the unbiasedness of the coefficients estimated under the different models; however, this does not mean the assumptions are realistic and convincing. By comparing the results from the different statistical models, I will examine which model could give us the most persuasive answer to causal question and under what assumptions in Chapter 5.

Chapter Ⅴ

RESULTS

5.1. Descriptive Statistics

The descriptive statistics were summarized in Table 8. Table 8 suggests that there were 956 boys, 37% of patients, and 1643 girls, 63% of patients, the median birth order was 2 with the interquartile ranging from 1 to 2, the mean birth weight was 2916.9g with ±449SD, the mean gestational age was 38.7 weeks with ±1.7 SD, the mean age of father was 31.9 years ±5.8 SD and the mean age of mother was 29.5 years old with ±4.6 SD. For the clinical characteristics, the mean age at the Kasai procedure was 67.6 days with ± 29 SD, 4% of patients was type I, 8% was type I-cyst, 2% was type II and 85% was type III, 5% of patients had associated anomalies, 42% of patients had cholangitis, usodeoxycholic acid (87%) and steroid 88% were frequently used. 1-year survival of native liver was achieved by 1551 patients (60%).

5.2. Bivariate Analysis

Table 11 shows the results of the bivariate analysis of 1-year native liver survival without jaundice by the timing of the Kasai procedure. The results show that there are

statistically significant differences for the Kasai procedure at less than 30 days (p=<0.001) and 45 to 60 days (p=0.001) with a larger percentage of patients in the favorable outcome group, while the Kasai procedure at 76 to 90 days (p=0.004), 106 to 120 days (p=0.002) and over 121 (p<0.001) had a larger percentage of patients in the poorer outcome group.

Figure 8 shows the 1-year native liver survival without jaundice and the age at the Kasai procedure. There is a descending trend in the relationship between 1-year native liver survival without jaundice and the age at the Kasai procedure; however, the relationship does not appear linear. The slope within the first 30 days appears steeper than the slope thereafter and the slope between 31 days to 55 days appears to almost flatten.

Figure 9 shows the relationship between the 1-year native liver survival without jaundice and the Kasai procedure more specifically across by two groups based on the timing of the Kasai procedure at less than 60 days and 60-120 days. The relationships are consistent with the results of the bivariate analysis showing almost no slope for the Kasai procedure done at age 31 to 45 days and those done between 61 to 75 days. The different timing of the Kasai procedure seemed to have different magnitudes of the effect on 1-year native liver survival without jaundice.

Figure 10 shows a slightly concave relationship between 1-year survival of native liver and the Kasai procedure performed between 31 to 70 days of age. This simple image

suggests that age of the Kasai procedure between 30 and 50 days has a slightly positive relation with the outcome.

Table 12 shows the bivariate analysis evaluating sociodemographic characteristics and clinical and hospital characteristics with relation to 1-year native liver survival without jaundice. Survival with native liver without jaundice did not appear to be associated with gender, birth order, birth weight, gestational age, and age of parents. The results were in accordance with the previous studies showing no association between the disorder and sociodemographics (Superina et al., 2011; The et al., 2007*)*.

The results showed statistically significant differences in types of obstruction, the use of dehydrocholic ($p<0.001$), and ursodeoxycholic acid ($p<0.001$), and hospitals with higher caseload ($p=0.01$) which had larger proportions among the native liver survival group. In contrast, associated anomalies and cholangitis had larger proportions among the non-native liver survival group.

5.3. Multivariate Analysis

Figure 11 shows the histogram of the timing of the Kasai procedure. It has two peaks existing at around 60 days and around 80 days. The mean age of the Kasai procedure was 67.6. Therefore, in the basic logit regression, the reference group was

defined as patients receiving the Kasai procedure between ages 61 to 75 days.

The results of logit/probit regressions without adjustment for covariates are presented in Table 13 in order to assure the comparability between two regressions with different link functions. Comparison to the reference group, receiving the Kasai procedure at less than 30 days of age increased the log odds of 1-year native liver survival without jaundice by 0.553 and it was statistically significant at ($p<0.05$) By taking the exponential of the log-odds ratio, I calculated an odds ratio of 1.7. This suggests that patients who received the Kasai procedure less than 30 days of age were 1.7 times more likely to achieve 1-year native liver survival without jaundice. Interestingly, age 31 to 45 and 46 to 60 days at the Kasai procedure were not statistically significant, though their coefficients suggested a positive relationship with the outcome. These results were consistent with observations exhibited in Figure 9 that showed age of the Kasai procedure between roughly around 30 and 55 days having no relationship with the 1-year native liver survival without jaundice. Their magnitudes are relatively small (31-45 days: 0.0153; 46-60 days: 0.127).

Unlike the results for the earlier timing of the Kasai procedure, the coefficients of the four categories corresponding to later the Kasai procedure showed negative associations with the outcome. Both the magnitudes and statistical significance became larger with later timing of the Kasai procedure except for the group having the Kasai

procedure between 91 to 105 days. Those who underwent the Kasai procedure after 120 days of age showed decreased log odds for 1-year native liver survival without jaundice by 1.3. This equated to them being 0.27 times as likely to achieve 1-year native liver survival without jaundice.

For the probit regression (also shown in Table 13), the Kasai procedure at less than 30 days of age in comparison to the Kasai procedure at 61 to 75 days increased the z-score of 1-year native liver survival without jaundice by 0.335 and was statistically significant at the 5% level. The results of the probit regression were similar to the results of the logit regression with regards to the magnitude of association and statistical significance. In order to assure their comparability, the ratios of their coefficients were calculated and are shown in the last column of Table 13. The ratios of the coefficients between logit and probit appeared constant across the different categories; therefore, their results were comparable in terms of signs and magnitudes of the coefficients.

Table 14 described the covariates selection approach. Among the sociodemographic variables included in the analysis, only birth weight was found to meet the threshold criteria of *p*-value was less than 0.20., The types of obstruction, associated anomalies, cholangitis, and hospital caseload were included in the final model because their associations with the outcome are previously suggested and also statistically significant in

the bivariate analysis. Among the use of choleretic agent, both dehydrocholic and ursodeoxycholic acid were included because of prior suggestion and statistical significance in the bivariate analysis. Steroid was included because results from a previous RCT study (Bezerra et al., 2014) suggested its negative impact on the short-term prognosis of patients.

The results of the multivariate logit regressions with full covariates, selected covariates and with hospital fixed effect are shown in Table 15. The reference group was patients whose age at the Kasai procedure was between 61 to 75 days and who had an anatomic pattern of type III. As opposed to what was observed in the basic logit regression without any covariates, the Kasai procedure at less than 30 days of age was no longer statistically significant. However, the other earlier categories of the timing of the Kasai procedure tuned into not statistically significant as well, although their direction of effect suggested positive associations with 1-year native liver survival without jaundice.

In contrast, the coefficients of all of four later categories of the timing of the Kasai procedure showed negative impacts on the outcome. Both the magnitudes and statistical significance become larger with later timing of the Kasai procedure except for the group having the Kasai procedure at 91 to 105 days. For those who underwent the Kasai procedure after 120 days of age, there was a decrease in log odds by 1.3, suggesting an

association with 1-year native liver survival without jaundice. This can be interpreted as them being 0.27 times as likely to achieve 1-year native liver survival without jaundice holding all other variables constant.

Type I, type I with cyst and type II obstructions were all independently associated with increased probability of achieving 1-year native liver survival without jaundice as compared to the type III group. Of all patients, 4% (n=101) were type I, 8% (n=212) were type I with cyst, 2% (n=54) were type II, and 86% (n=2065) were type III. When compared to those who had type III obstruction, patients with type I obstruction had a log odds of 0.757 for 1-year native liver survival without jaundice. This represented a 2.13 times greater likelihood of achieving 1-year native liver survival without jaundice for type I obstruction patients compared to patients with type III obstruction independent of other covariates in the model. For type I with cyst, the log odds ratio was 1.487 and represented a 4.4 times greater likelihood of achieving the favorable outcome. Type II patients showed a log odds ratio of 0.76 with a magnitude of effect that was similar to type I, but with less statistical significance. These results confirmed that patients having the type III anatomic pattern had a lower probability of achieving 1-year survival of native liver.

Associated anomalies were negatively associated with the outcome and showed borderline statistical significant at the 5% level, and cholangitis showed a negative

relationship with the outcome after adjustment for covariates. The log odds ratio of cholangitis was -0.254 which was equivalent to odds ratio of 0.78.

Among the variables related to choleretic agents, the dehydrocholic and ursodeoxycholic acid showed positive associations with the outcome after adjustment for potential confounding. Particularly, the ursodeoxycholic acid showed a log odds ratio of 0.587, which indicated a 1.7 times greater likelihood of achieving 1-year survival of native liver compared to those who were not treated with this agent. In contrast, steroid use was independently associated with a decreased probability of achieving the favorable outcome showing a log odds of 0.413. Hospital caseload showed a suggestive association with a log odds ratio of 0.219 which is equivalent to an odds ratio of 1.24.

The two models in the analysis, one with full covariates and the other with selected covariates showed very similar results despite difference in sample size availability. A ROC was produced (Figures 12 and 13), and used as a measure of model fit. The areas under the ROC were similar for the two models; 0.659 for the full model and 0.656 for the selected covariates model. This suggested that the two models performed similarly with respect to predicting the outcome.

5.4. Stratified Multivariate Analysis

Figure 14 showed the relationship between the 1-year native liver survival without jaundice and the timing of the Kasai procedure by the types of obstruction. There were 91 subjects in the type I group, 199 in type I with cyst, 47 in type II, and 2,065 in type III. Surprisingly, for the type I and type II groups, there were no patients who had a Kasai procedure after 200 days of age. The four groups showed very different trends in the relationship between the timing of the Kasai procedure and the outcome.

Table 16 showed the predictors of 1-year native liver survival without jaundice stratified by the types of obstruction adjusted for sociodemographics, clinical and hospital characteristics. The reference group was patients whose age at the Kasai procedure was between 61 and 75 days. The coefficients associated with the early timing of the Kasai procedure compared to the reference group were not statistically significant, with the exception of the type III patients who had a Kasai procedure at less than 30 days when evaluated at the 10% significance level. This finding indicates that the Kasai procedure at less than 30 days of age in comparison to the Kasai procedure from 61 to 75 days increased the log odds of 1-year native liver survival without jaundice by 0.553 and it was statistically significant at the 5% level. For type III patients, the result showed a log odds ratio of 0.43 with borderline statistical significance for patients who had the Kasai procedure at less than

30 days, which was similar to the result of the multivariate logit regression (odds ratio of 0.32. z=1.49). However, in the model stratified by types of obstruction and controlling for covariates, not only were the estimates no longer statistically significant for timing of the Kasai procedure less than 30 days of age.

Generally, the coefficients of most of the four later categories of the timing of Kasai procedure showed a negative impact on the outcome regardless of their type of obstruction. For the type I group, timing of the Kasai procedure was not independently associated with the outcome after controlling for potential confounding. For the type I with cyst group, only the timing of the Kasai procedure over 121 days showed a decreased probability of achieving 1-year native liver survival without jaundice. For the type II group, though affected by small sample size, the timing of the Kasai procedure from 76 to 90 days and over 121 days showed a negative association with the outcome. Both the magnitude and statistical significance increased with later timing of the Kasai procedure for the group receiving the Kasai procedure at 91 to 105 days. In type II group, receiving the Kasai procedure after 121 days of age decreased the log odds of 1-year native liver survival without jaundice by 1.3. This means that they were 0.27 times as likely to achieve 1-year native liver survival without jaundice independent of other covariates in the model.

For the type III group, both the magnitude and statistical significance increased

with later timing of the Kasai procedure, except for the group having the Kasai procedure at 91 to 105 days. Receiving the Kasai procedure at 106 to 121 days of age decreased the log odds of 1-year native liver survival without jaundice by 0.83. This means that these patients were 0.43 times as likely to achieve 1-year native liver survival without jaundice holding other covariates constant. The Kasai procedure after 121 days of age also decreased the log odds of 1-year native liver survival without jaundice by 1.17, meaning that these patients were 0.3 times as likely to achieve 1-year native liver survival without jaundice holding other covariates constant.

The birth weight and associated anomalies were not associated with the outcome. For type I patients, patients with cholangitis showed to be 0.14 times as likely to achieve 1-year native liver survival without jaundice after adjustment for potential confounders. Among patients with the type I with cyst, results for cholangitis suggests to be 0.22 times as likely to achieve 1-year survival of native liver after adjustment. The result for patients with type III was also in the same direction showing them to be 0.79 times as likely to achieve 1-year native liver survival without jaundice. These results for the effect of cholangitis across the subgroups suggest that patients who are type I may be more likely to be affected by the negative impact of cholangitis on 1-year native liver survival without jaundice followed by type I cyst and type III. The relationship in type II patients could not be evaluated due to

limitations in sample size.

The results on the use of choleretic agent suggest that dehydrocholic was not associated with the outcome. The ursodeoxycholic acid was statistically significant in the type I, type I with cyst, and type III groups with the log odds ratios of 4.27, 1.61, and 0.54, respectively. Steroid treatment was negatively associated with the outcome in all four groups but statistically significant only in the type I group with the log odds ratio of -4.86. The statistically significant effect of the higher hospital case load was observed only in the type III group with a log odds ratio of 0.25.

5.4. Fixed Effect

In regard to the signs and magnitudes of the effect of each variable, the results obtained for the FE were very similar to the results on the multivariate logit regression with selected covariates except for the significance of the timing of the Kasai procedure less than 30 days. In the FE model, it turned out to be statistically significant at 5% level. It showed that increased log odds by 0.458 and was associated with 1-year native liver survival without jaundice. This can be interpreted as them being 1.6 times as likely to achieve 1-year native liver survival without jaundice holding all other variables constant.

In contrast, the coefficients of all of four later categories of the timing of the Kasai

procedure showed negative impacts on the outcome. Both the magnitudes and statistical significance became larger with later timing of the Kasai procedure except for the group having the Kasai procedure at 91 to 105 days.

5.5. Instrumental Variable

Two possible instrumental variables were considered for the analysis: 1) whether or not patients received any kind of screening prior to the Kasai procedure and 2) length of stay at last hospital admission prior to having the Kasai procedure. For the first instrumental variable, there was a variable describing whether or not patients received any kind of screening prior to being referred to the hospital for detail examinations.

A total of 23 patients received some form of screening prior to a Kasai procedure showed in Table 10, representing 1% of the total subgroup patients. The other screening methods included measurement of conjugated bile acids in dried blood spots using tandem mass spectrometry, serum conjugated bilirubin after birth, measuring urinary sulfated bile acid, and measurement of fecal conjugated bilirubin by near-infrared reflectance spectroscopy and each of them was performed in the selected hospitals and was offered partly as clinical researches.

The limitation of using the screening variable as the instrument was the fact that

the number of observation was too small; therefore, it is not sufficient to explain any variations in the timing of the Kasai procedure. Since 2012, the national stool color card has been implemented, so in few years the larger numbers of patients who have screening are expected and makes IV analysis possible with this instrument.

The second instrument was the length of stay at last hospital admission prior to having the Kasai procedure. It is used as a proxy for the accessibility to the nearest hospital with high-level hospital which means that the hospital with a specialized surgeon and facility.

In the IV analysis, I dichotomized the timing of the Kasai procedure less or more than 30 days, 45 days, 60days, 75 days, 90 days, 105 days and 120 days for simplicity of interpretation and discussion.

The results of 2SLS are summarized in Table 18. The results on OLS and probit regression without any adjustments are shown in the first two columns, adjusted for sociodemographics, hospital and clinical characteristics are also shown in the third and fourth columns of Table 18 as baseline references. The results in both OLS and probit regression suggested that the timing of the Kasai procedure was independently associated with the 1-year native liver survival without jaundice while any of the timings of the Kasai procedure in 2SLS do not significantly explain the probability of 1-year native liver survival without jaundice. The signs and t-value or z-score were consistent in two models which

confirmed that the choice of the model did not affect the results in the analysis.

5.6. Discussion

Recalling $\log(p/(1-p))SNL_{kt} = \alpha_0 + \beta_1 K_{kt} + \beta_2 X_{kt} + \alpha_{kt} + \varepsilon_{kt}$ --- (7), in this empirical analysis, the outcome was measured as 1-year native liver survival without jaundice. The timing of the Kasai procedure was divided into 8 groups ranging from the group having the Kasai procedure at less than 30 days incremented by 15 days up to over 121 days because some of the previous studies suggested that earlier intervention was associated with better prognosis (Schreiber et al., 2007; Chardot et al 1999, 2013). While the traditional approach for defining the timing of the Kasai procedure was to divide it into two groups with a cut-off point at 60 days of age, the categorization into 8 groups used in this research allowed us to see not only the effect of the timing of the Kasai procedure before 60 days, but also after 90 days, which has previously been reported to be associated with worse prognosis (Chardot et al., 1999, 2001).

The full covariates included sociodemographics and clinical and hospital characteristics. The incidence of associated anomalies seemed low in this study, and this was likely because that this study did not include cardiac anomalies. Previous studies have suggested the incidence as 10 to 15 percent of infants with BA (Schreiber et al., 2007) and have shown to be related to poorer outcomes. The incidence of cholangitis in this study

seems somewhat low (42%) compared to other previous studies that ranged from 40% to 90%, with the majority of patients experiencing at least one episode prior to two years of age (Luo and Zheng, 2008; Sokol et al., 2007).

As previous literature (Muise et al., 2006; Nio and Muraji, 2013; Nio et al., 2006; Nio et al., 2010; Stringer et al., 2007; Superina et al., 2011) and the figures from this study also suggested, the types of obstruction have a substantial impact on patient outcomes. The results from the multivariate analysis adjusted by types of obstruction showed that receiving the Kasai procedure within less than 30 days of birth had a suggestive impact (z=1.79) only for the type III patients. According to this result, there was a possibility that a Kasai procedure at less than 30 days might improve the short-term outcome for patients who have a type III obstruction. This difference might be explained by the fact that types of obstruction may be associated with the baseline severity and progress of the disease as it was previously reported that Type I BA (atresia of the common bile duct), or so called correctable atresia, showed a better prognosis when comparing to type II and III (Superina et al., 2011).

The small sample size (particularly for type II [n=49]) might explain the lack of statistical power, and therefore, the lack of significance of the coefficients. Also the confidence intervals were rather large for type I, type I with cyst, and type II. However, as observed both in Table 16 and Figure 14, the signs, magnitudes, and general trends

between the timing of Kasai and 1-year native liver survival were different among four different types of obstruction. For type I, type I with cyst and type II, although not statistically significant, the earlier timing (less than 45 days) of Kasai procedure had a negative impact on achieving 1-year native liver survival in general, and this was associated with a contraction of the effect for type III. These findings were also depicted from Figure 14, suggesting that there were almost flat relationships between the earlier timing of Kasai procedure and the outcomes for type I and type I with cyst. From these observations, for type I and type I with cyst, there seemed to be a different effect of the timing of the Kasai procedure on the outcome, with the opposite directional impact. For type II, a larger sample size would be required to reach a meaningful conclusion.

The hypothesis behind the application of FE in the study was that the timing of the Kasai procedure varied within hospitals but any hospital characteristics affecting the timing of the Kasai procedure and outcome did not vary within each hospital. In this setting, within hospital correlation of the timing of the Kasai procedure and the outcome should not be detected. My hypothesis in FE regression was that if the variable for hospital characteristics in the multivariate logit regression (defined as hospitals with higher case load compared to the mean caseloads out of all hospitals) accounted for the variations in hospital characteristics, then the results from the multivariate logit regression and FE would be the

same.

However, the results from the FE suggested a slightly different effect for the timing of a Kasai procedure at less than 30 days. In the FE model, it turned out to be statistically significant ($p<0.05$). FE results showed an increased log odds of 0.458 and were associated with 1-year native liver survival without jaundice, suggesting that receiving the Kasai procedure at less than 30 days after birth was 1.6 times as likely to achieve this outcome independent of other covariates in the model. Therefore, there still might be unmeasured hospital characteristics that would affect both the timing of the Kasai procedure and the outcome.

The limitation of FE is that it can only exclude any impact associated with time-invariant within-hospital characteristics that may affect the dependent variable. The FE analysis did not account for variations in innate child healthiness or random shocks to timing of the Kasai procedure among patients within the same hospital.

For the IV analysis, I used a shorter length of stay at the last hospital admission where patients had their Kasai procedure as a proxy for the accessibility to the nearest high-level hospital. My hypothesis was that the longer the length of hospital stays prior to the Kasai procedure, the higher the probability that patients may live farther from a high-level hospital. The rationale was that if patients lived close to a high-level hospital then the

screening could be done on an outpatient basis, whereas if patients lived far from the hospital, they would prefer to be examined and screened on an inpatient basis.

The first stage results summarized in Table 17 suggested that in the first four groups, having the Kasai procedure less than 30 days, 45 days, 60 days, and 75 days, the coefficients of length of stay at last hospital admission prior to receiving the Kasai procedure were statistically significant. These findings suggest that the IV was related to the timing of the Kasai procedure while holding other predictors constant.

To see the strength of the instrument, I performed the diagnostic test. The most common diagnostic test of the first-stage is the F - statistic on the excluded instrument. A rule of thumb is that the F should be above 10 (Stock, Wright and Yogo, 2002). F-statistics above 10 were observed in KP less than 30 days, 45 days, 60 days, and 75 days (10.7, 42.6, 74.4, 15.9 accordingly) but not for the rest of group. These results suggested that the first four IVs were strong, but not for the later timing of the Kasai procedure. The IV used for the timing of the Kasai procedure after 90 days were weak; therefore, the IV estimates may had have high variance, and their partial first-stage F statistics were less than 10, therefore, it was better to use inference methods other than 2SLS.

Table 19 and 20 show the full results on adjusted OLS and probit regression and Table 21 and 22 show the results on 2SLS and probit with IV. While in OLS and probit

regressions, type I, type I with cyst, and type II anatomic patterns showed a positive association with the outcome, as did use of ursodeoxycholic acid and the hospital with higher case load. However, type III anatomic pattern, cholangitis and use of steroid therapy showed a negative impact on the outcome. In the results of the 2SLS and probit with IV models, these variables stayed as significant as they were in the OLS and probit regression while the timing of the Kasai procedure was found to be statistically nonsignificant. These findings suggested that types of obstruction, cholangitis, the use of choleretic agents, and the hospital caseload were mediating variables that interacted with the timing of the Kasai procedure in determining 1-year survival of native liver with jaundice.

The difference in the significance of the effect of the timing of the Kasai procedure in OLS and 2SLS leading to wonder which models would be most accurate. Therefore, I performed the Wu-Hausman test for the 2SLS model; the null hypothesis was that there was no difference between the OLS and the 2SLS model. The hypothesis had assumed that there had not been an endogeneity problem. Moreover, if there was no endogeneity problem, both OLS and 2SLS would be consistent, but 2SLS was inefficient. Estimating the 2SLS model showed that the test was not significant for the timing of the Kasai procedure (KP<30 p=0.65, KP<45 p=0.77, KP<60 p=0.83, KP<75 p=0.94, KP<90 p=0.69, KP<105 p=0.68, KP<120 p=0.61). The Wu-Hausman tests were not statistically significant, thus

failing to reject the null hypothesis that the estimates of both the models were consistent. This suggested that the 2SLS model was not the preferred model and the timing of the Kasai procedure was not endogenous, and therefore, IV was not the preferred estimator. Also the results of the analyses using OLS and probit models were not radically different and that provided us with some assurance that the choice of model was not critical to my results.

Chapter VI

CONCLUSIONS

6.1. Summary of Findings

The goal of this study was to use different statistical models to assess the associations and to answer causal question on the effect of the timing of the Kasai procedure on patient outcomes. To date, none of the available literature on the timing of the Kasai procedure's effect on BA patient outcomes has addressed the problem of endogeneity. Moreover, only three studies using population-based datasets tried to control for confounding (Lien et al., 2011; Chardot et al.,1999; Chardot et al., 2013). Their results suggested that in addition to the timing of the Kasai procedure, the anatomic pattern of BA, existence of associated anomalies, prophylactic antibiotics, and hospitals with high caseload had independent association with patient outcomes.

I reviewed the literature in economics and education specifically to examine the effect of health status on education to learn the methodology and approaches used to mitigate potential biases. Four approaches that were reviewed include DiD, PSM, FE, and IV approach. Among the four approaches that were reviewed, multivariate logit/probit regression, hospital FE, and IV were selected for use in this study.

Figure 15 shows the multiple estimation results with 95% confidence interval combined. It displays four multivariate logit regression models including with no covariates (blue), with full covariates (red), with selected covariates (orange), and with hospital fixed effect (green) that examined the effect of Kasai procedure on 1-year native liver survival. It presents the four models by plotting parallel lines for each of them grouped by coefficients. As it is suggested in Figure 15, in terms of the magnitude of the coefficients, the types of obstruction have the largest impact on the outcome though the confidence intervals are rather large. Particularly, there was a difference in terms of the odds ratios between the blue dot and the rest of the dots for the timing of Kasai procedure less than 30 days. The blue dot is based on the bivariate analysis of 1-year native liver survival without jaundice by the timing of the Kasai procedure less than 30 days. There was a statistically significant difference for the Kasai procedure at less than 30 days (p=<0.001) with a larger percentage of patients in the favorable outcome group. However, when sociodemographic, clinical and hospital characteristics were controlled, the log odds ratio of achieving the 1-year native liver survival without jaundice for the Kasai procedure at less than 30 days of age decreased and became statistically non-significant in all but the model with the hospital fixed effect (green dot). This suggests that the effect of Kasai procedure at less than 30 days was more susceptible to other factors affecting the outcome such as the types of obstruction and

hospital characteristics.

The results from this study differed from those of previous studies in that the findings showed that patients who underwent a Kasai procedure at ages 31 to 45 days and 46 to 60 days did not have a statistically significant improvement in the outcome variable, though their coefficients did suggest a positive relationship with the outcome. These results were consistent with observations presented in Figure 8 that showed age of the Kasai procedure between 30 and 55 days having no relationship with native liver survival without jaundice and their magnitudes effect are relatively small (31-45 days: 0.0153; 46-60 days: 0.127) while holding everything else constant.

While previous studies and also the bivariate analyses showed that there was a negative association between the timing of the Kasai procedure and patient outcomes, the analysis of the timing of the Kasai procedure indicated that as compared to the reference group defined as patients receiving the Kasai procedure between ages 61 to 75 days, the effect of earlier timing of the Kasai procedure that were at less than 30 days, 31-45 days, and 46-60 days, did not significantly explain 1-year native liver survival without jaundice after adjusting for covariates. The stratified logit regression by types of obstruction suggest that the timing of the Kasai procedure did not affect the outcome in the patients with type I and type I with cyst. Only in the type III group did the timing of the Kasai procedure at less

than 30 days showed a suggestive association to the increased probability of achieving 1-year native liver survival. Later timing of the Kasai procedure was associated with decreased probability of achieving 1-year native liver survival except for the Kasai procedure at 91 to 105 days.

As previously suggested, the type I, type I with cyst and type II obstructions were all independently associated with increased probability of achieving 1-year native liver survival without jaundice as reference to the type III group. These results also confirmed that patients having the anatomic pattern of type III have a lower probability of achieving 1-year survival of native liver.

Cholangitis showed a negative relationship with the outcome after adjustment for covariates as suggested in the previous study. However, associated anomalies were not statistically significant in the multivariate logit regressions. This contradicts with the findings from previous studies but the difference may have been attributed to the definition and classification of the associated anomalies in each study.

Among the variables related to choleretic agents, only ursodeoxycholic acid showed the positive associations with the outcome after adjustment for potential confounding. In contrast, steroid therapy was independently associated with a decreased probability of achieving the favorable outcome.

Hospital caseload showed an association with the outcome while adjusting for covariates. While it showed a positive relationship to the outcome in the type III group (log-odds ratio of 0.23, $p<0.05$), it did not show the association to the outcome in the groups with type I or type I with cyst. Further analysis with hospital FE was performed to see if there were any hospital characteristics rather than the caseload effect that would affect the timing of the Kasai procedure and the outcome. The result on FE suggested the slightly different effect of the timing of the Kasai procedure less than 30 days. This suggests that there still might be unmeasured hospital characteristics that would affect both the timing of the Kasai procedure and the outcome.

By comparing the results on adjusted OLS and 2SLS, it revealed that types of obstruction, cholangitis, the use of choleretic agents, and the hospital caseload were mediating variables that interacted with the timing of the Kasai procedure in determining 1-year survival of native liver with jaundice.

In sum, as opposed to the conventional wisdom, the results from the multivariate logit regression adjusted for covariates suggested that for patients with the obstruction types I and I with cyst, the timing of the Kasai procedure was not associated with 1-year native liver survival. As previously suggested, cholangitis and use of steroid showed negative associations to the outcome after adjustment for covariates. However, associated

anomalies were not statistically significant in the multivariate logit regressions which contradict with the previous findings but may be attributed by the definition and classification of the associated anomalies in each study. Ursodeoxycholic acid showed a positive association with the outcome. Hospital caseload also showed a positive association with the outcome while adjusting for covariates.

6.2. Limitations: Dataset, Assumptions, and Interpretation

The limitation of this study is largely driven by the nature of this rare disease of which its etiology is not clear. There are both observed and unobserved processes that lead to determine the timing of the Kasai procedure, and these factors are correlated with survival with native liver, either directly or indirectly but very little is known about how these factors interrelate to each other.

The dataset had its strength in terms of the largest subjects and was population-based. However, it also had some limitations. The outcome was not observed or was unknown for 122 patients, and was therefore excluded from the study.

A total of 16 patients were also excluded from the study because they did not receive the Kasai procedure. While the outcomes for these 122 patients were unknown, 6 out of 16 (37%) had died within one year. However, the clinical outcomes for the rest of 10

patients were not known since the JABR dataset did not contain information on liver transplantation for these patients. It is possible that they might have gone directly to the transplantation as a result of late diagnosis.

Generalizability of the study results to the target population could be affected if the excluded patients were somehow different from those included in the study. In fact, for patients who underwent a Kasai procedure, the death rate was 6.8% within one year (117 out of 2605). I suspect that those patients who did not undergo a Kasai procedure had a different background which therefore might have affected the outcome.

Random attrition and missing completely at random would not affect the point estimates, but would have an impact on the standard errors associated with them (i.e. they would become larger). However, if patients who have been dropped out of the sample had lower or higher probability of survival of native liver than patients who remained in the sample, then the coefficients associated with the determinants would be biased.

The external validity of this study is expected to be high due to its population-based sample from a dataset that likely covers approximately 96% of BA patients in Japan. The Internal validity of this study is based on a set of assumptions made in the empirical analyses which were derived from two perspectives. The first perspective was that whether there were no unmeasured characteristics in the model was assumed. The multivariate

logit/probit regression and FE assumed that there were no unobservable characteristics that changed over time and can affect patient outcomes in the ways that were associated with selection and endogeneity biases. Potential problems associated with child's innate healthiness and hospital characteristics were ruled out in the analyses by controlling for birth weight and the indication of hospital with high caseload in the model or by FE. In addition to these, the analyses assumed no heterogeneity in timing of the Kasai procedure effects.

The second perspective was that when there were unmeasured characteristics in the model affecting both the timing of the Kasai procedure and patient outcome, and then the IV analysis should be used.

I used length of stay at the last hospital prior to the Kasai procedure as a proxy for the accessibility to the nearest high-level hospital in IV analysis. The diagnostic test to see the strength of the instrument suggested that IVs were not weak in the Kasai procedure at less than 30 days, 45 days, 60 days, and 75 days, but not for the rest of later timing of groups. Wu-Hausman test confirmed that the estimates of both the models were consistent. This suggested that the 2SLS model was not the preferred model and suggested that the timing of the Kasai procedure was not endogenous.

6.3. Further Research

The results should be analyzed in the long-term outcomes. Due to limited data access, the study only investigated thethe short-term outcome of 1-year native liver survival without jaundice. However, the results on short-term outcome may not necessarily be an ideal proxy for long-term outcomes. Ultimately, what matters to patients is not the short-term prognosis, but rather how to live with the disease and how it affects their lives in the long-run. In the JBAR dataset, a long-term outcome measurement such as 10- 20-year native survival is also available but not ready for the current analysis so in the near future, such data will be also used to assess the longer-term effect of Kasai procedure.

As suggested in the study findings and discussion, the effect of the timing of Kasai procedure differed for the different types of obstruction. Further research is warranted focusing on analyses using more interaction terms to see whether there is the presence of a significant interaction indicating the effect of the timing of Kasai procedure on the response variable that is different depending on the values of the other predictor variables.

I applied only one instrument in this study and the average treatment effects for the particular subgroups of the study may not generalizable to the entire subjects in the study. In order to further verify the validity and robustness of the results, I would have to apply more instruments within the study in order to compare the findings from the different IV analyses.

This would allow examining whether the IV analyses are aligned in both directions and magnitudes.

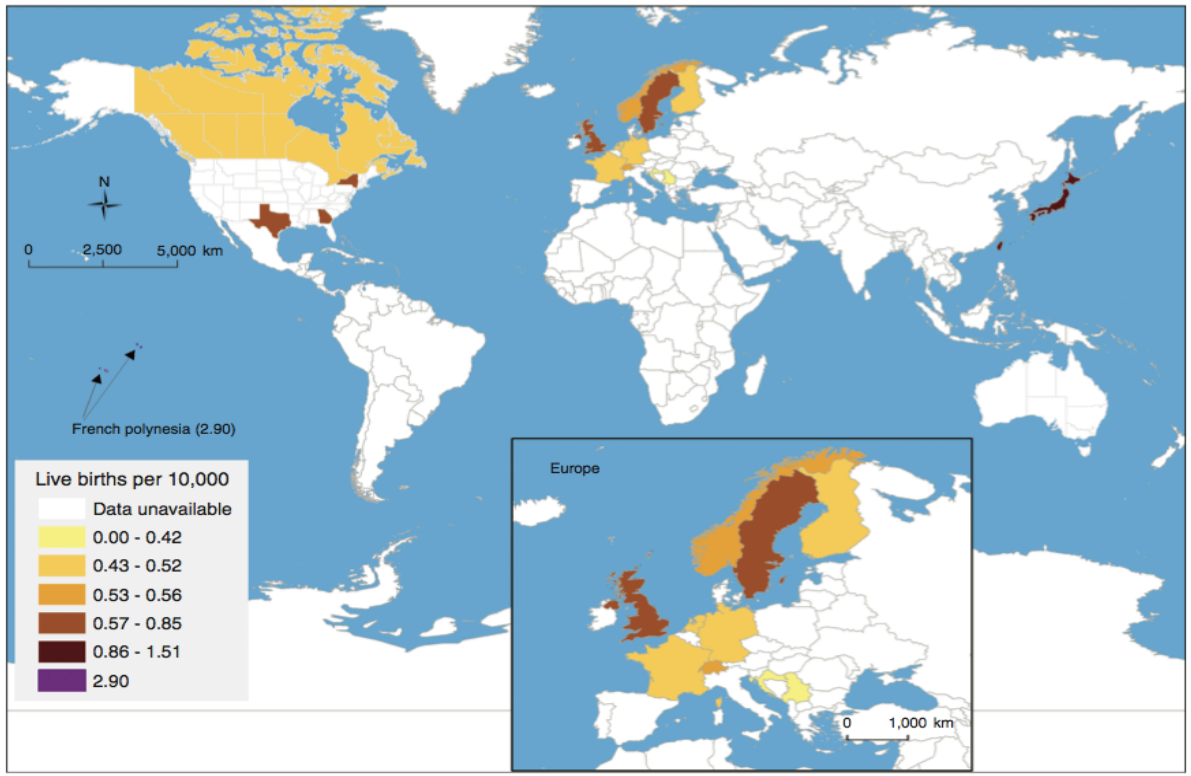
As discussed in Chapter 3 and 4, the recent government policy on the national stool color card implemetation launched in 2012 in Japan would be an ideal candidate to examine the causal effect of the timing of Kasai procesure. The current nationwide implementation of stool color card should provide clear insights regarding target patients and the ideal timing of the Kasai procedure. Although the purpose of the stool color card screening is to achieve the early surgical intervention, it is still not clear how early is early enough to achieve optimal patient outcomes. Policy makers should be clear on the goals and targets of the national stool color card implementation.

Cost-effective analysis among different screening modalities is also warranted. Distributing a stool color card might be less expensive than taking and testing clinical samples such as blood or urine. However, differences in the sensitivity of screenings could result in variance across subsequent outcomes, leading to variation in overall efficiencies.

It is my hope that international collaboration on pooling existing data on rare disease will occur in the near future which will answers to many unanswered questions and cast new hopes for patients with rare diseases and their families.

FIGURES AND TABLES

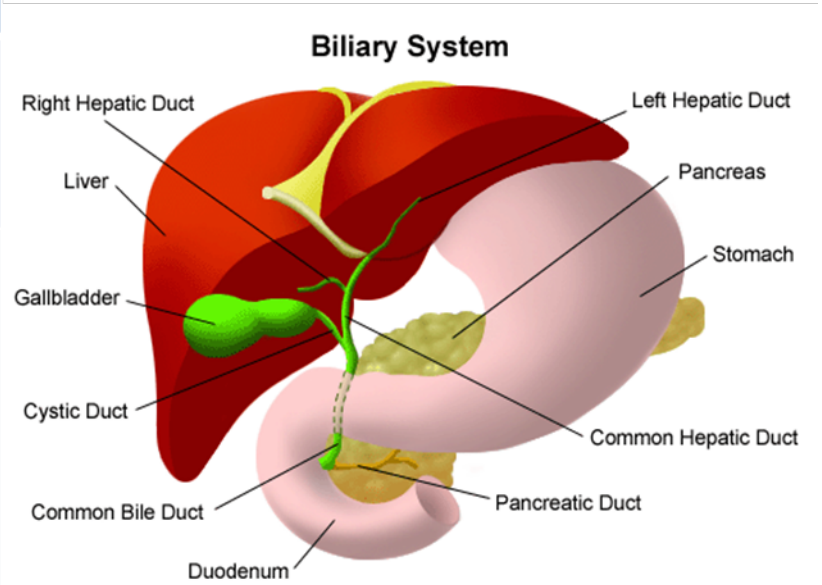
Figure 1 Incidence Rate of Biliary Atresia in the World



Source: Jimenes-Rivera et al (2013)

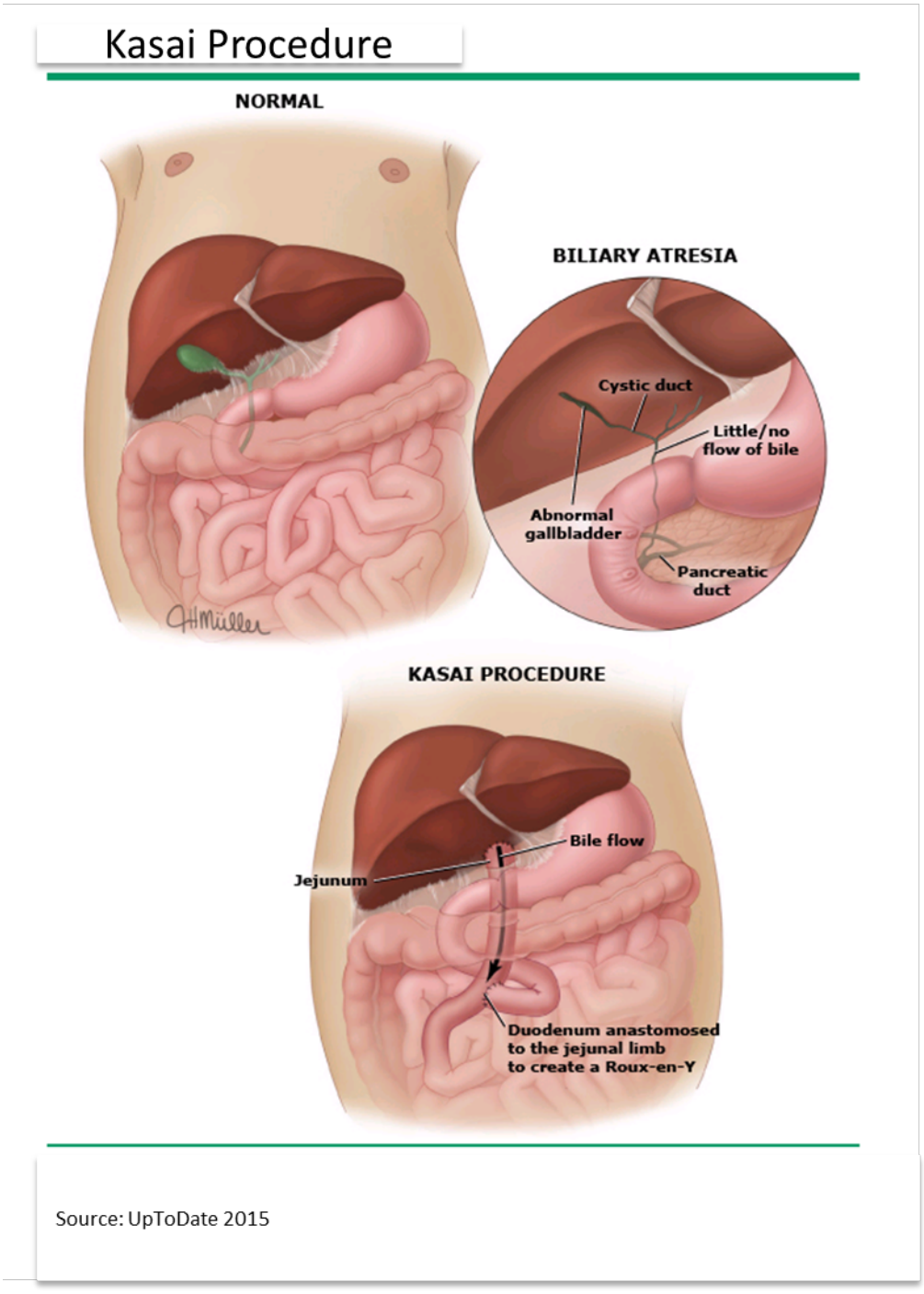
Figure 2 Anatomic Pattern of BA

Type	Definition
I	Level of obstruction is within the common bile duct. *Gallbladder contains bile.*
II	Level is within the common hepatic duct. *The gallbladder will not contain bile but a transection of the proximal remnant should show two distinct bile-containing lumens.*
III	Obstruction is within the porta hepatis. *No visible bile-containing proximal lumen.*



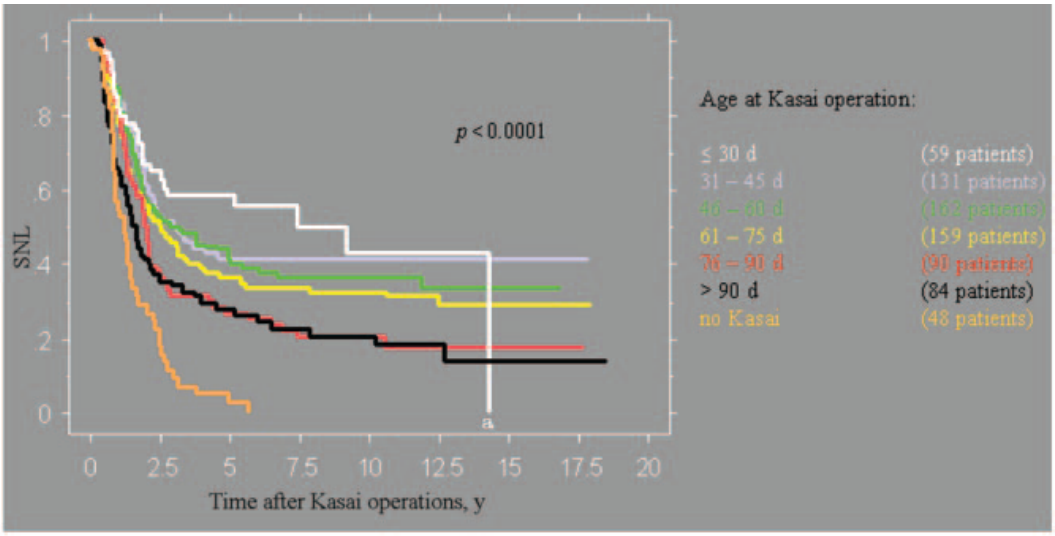
Source: Japan Biliary Atresia Society

Figure 3 Kasai Procedure



Source: UpToDate 2015

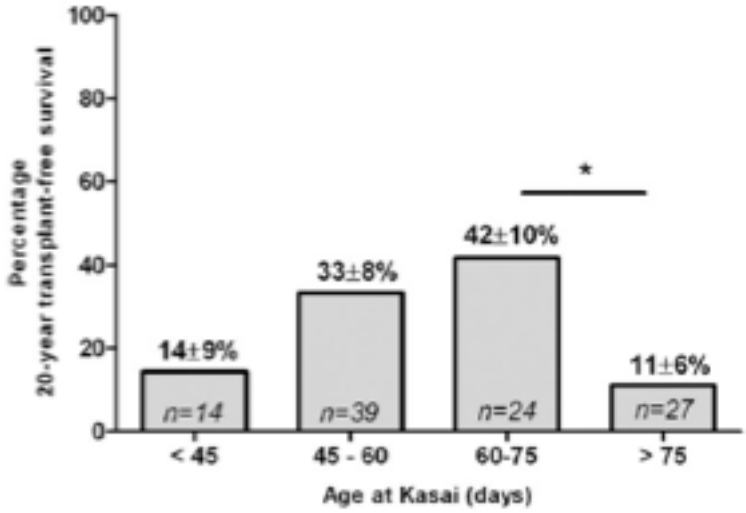
Figure 4 Kaplan-Meier and the log-rank test of Survival of Native Liver at 2-, 5-, 10-, and 15 years in France (2009)



Age at Kasai operation	*N* complete data sets for analysis	2-y SNL [SE] (*n* patients alive with native liver at 2 y)	5-y SNL [SE] (*n* patients alive with native liver at 5 y)	10-y SNL [SE] (*n* patients alive with native liver at 10 y)	15-y SNL [SE] (*n* patients alive with native liver at 15 y)
≤30 d	59	66.2% [6.3] (35)	58.1% [6.8] (21)	42.5% [9.6] (6)	— — (0)
31–45 d	131	65.5% [4.2] (79)	40.5% [4.5] (37)	40.5% [4.5] (20)	40.5% [4.5] (4)
46–60 d	162	57.8% [3.9] (89)	42.4% [4.1] (45)	36.1% [4.2] (18)	33.3% [4.7] (2)
61–75 d	159	57.1% [4.0] (82)	36.1% [4.1] (39)	32.3% [4.1] (24)	28.7% [4.3] (7)
76–90 d	90	52.4% [5.4] (43)	26.4% [5.0] (17)	19.5% [4.7] (8)	16.7% [4.8] (6)
>90 d	84	42.0% [5.6] (31)	27.3% [5.3] (15)	20% [5.0] (9)	13.4% [5.3] (2)

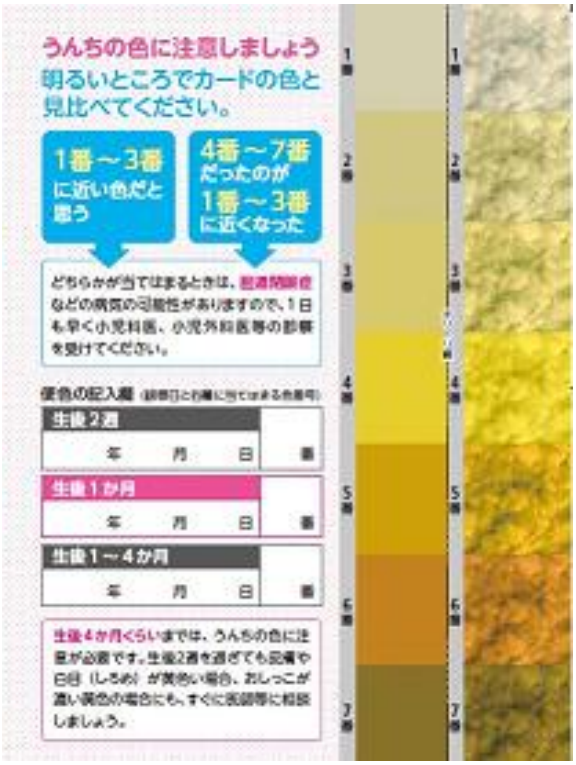
Source: Serinet et al (2009)

Figure 5 20-year transplant free survival of BA patients and timing of Kasai procedure in Netherland



Source: De Vries et al (2012)

Figure 6 Stool Color Card Chart



Source: Mother and Child Health Handbook (Japan)

Figure 7 Causal Diagram on Timing of Kasai Procedure and Survival of Native Liver

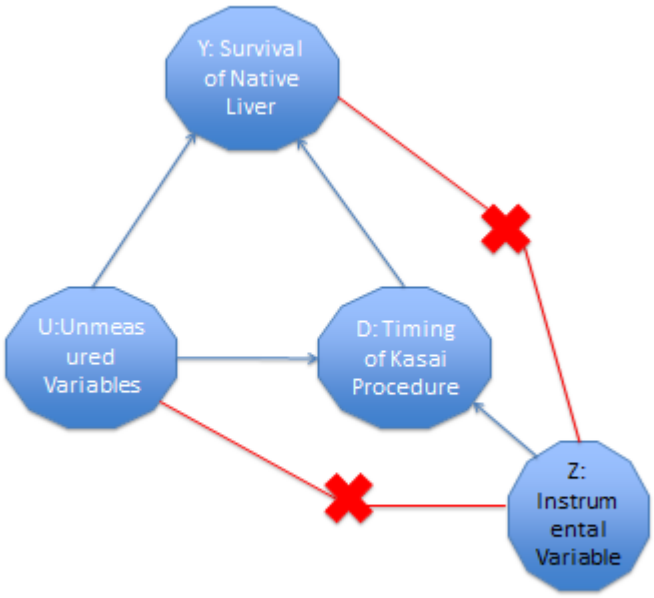


Figure 8 1-Year Native Liver Survival without Jaundice and Age at Kasai

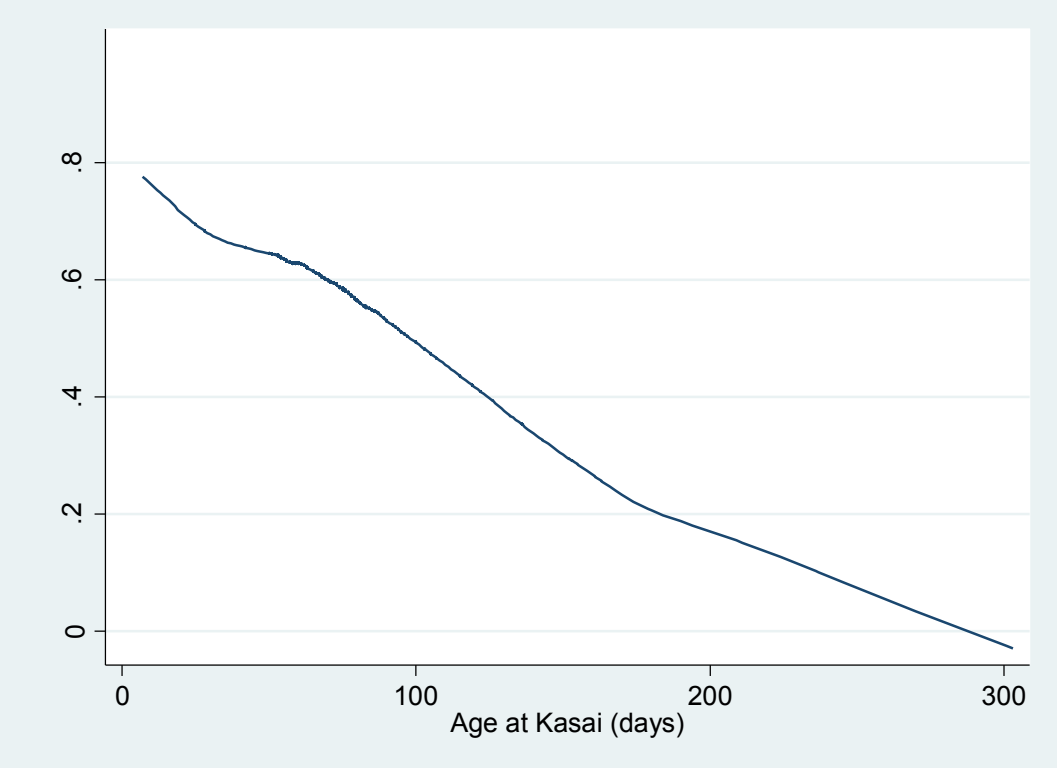


Figure 9 1-Year Survival of Native Liver without Jaundice and Age at Kasai less day (<60 days v.s. 60-120 days)

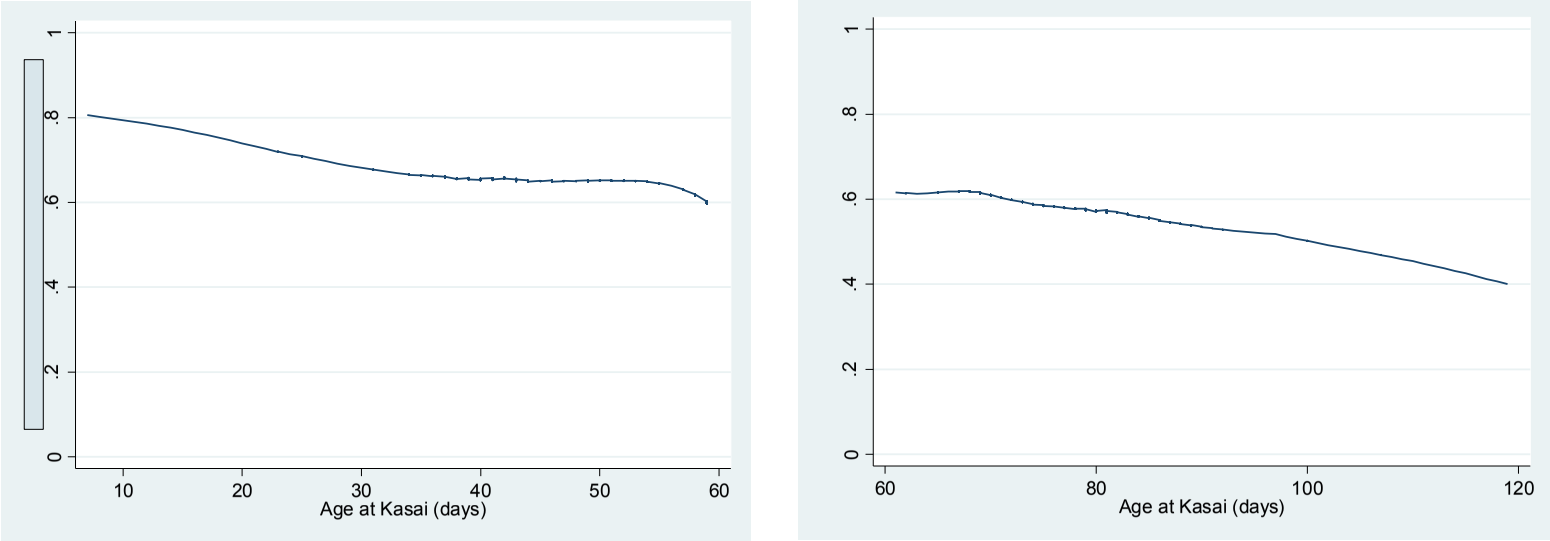


Figure 10 1-Year Survival of Native Liver without Jaundice and Age at Kasai less day (30-80 days)

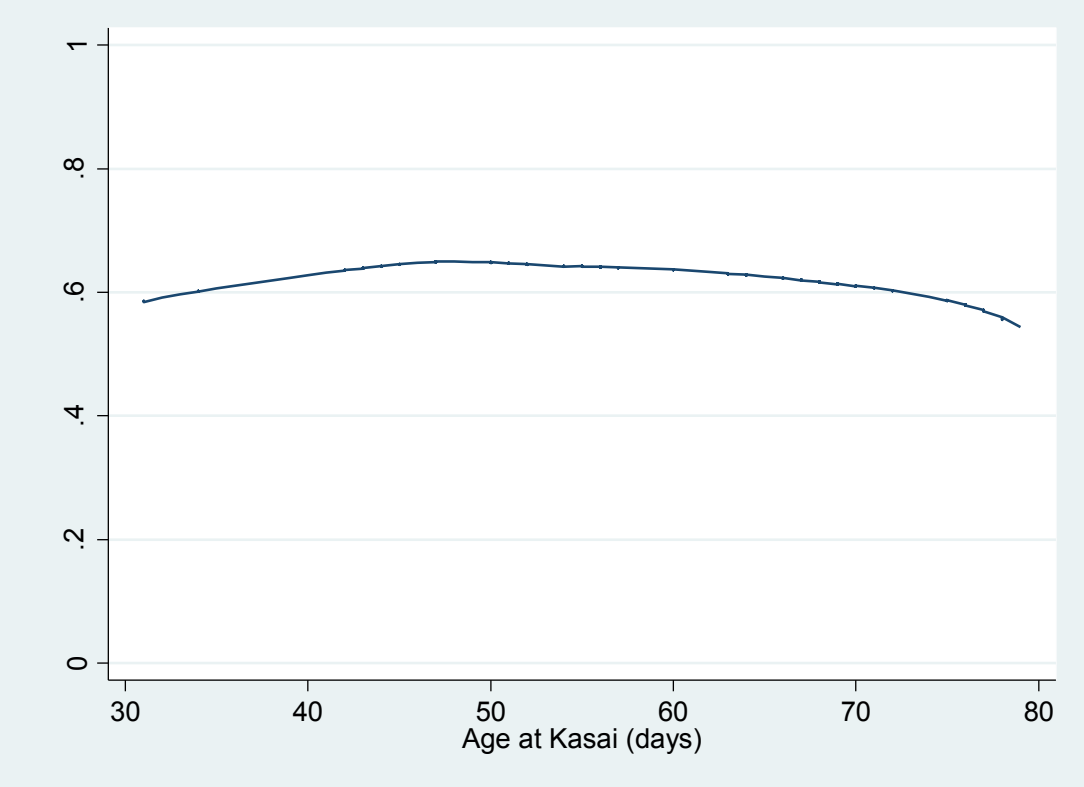


Figure 11 Histogram of Timing of Kasai Procedure (days)

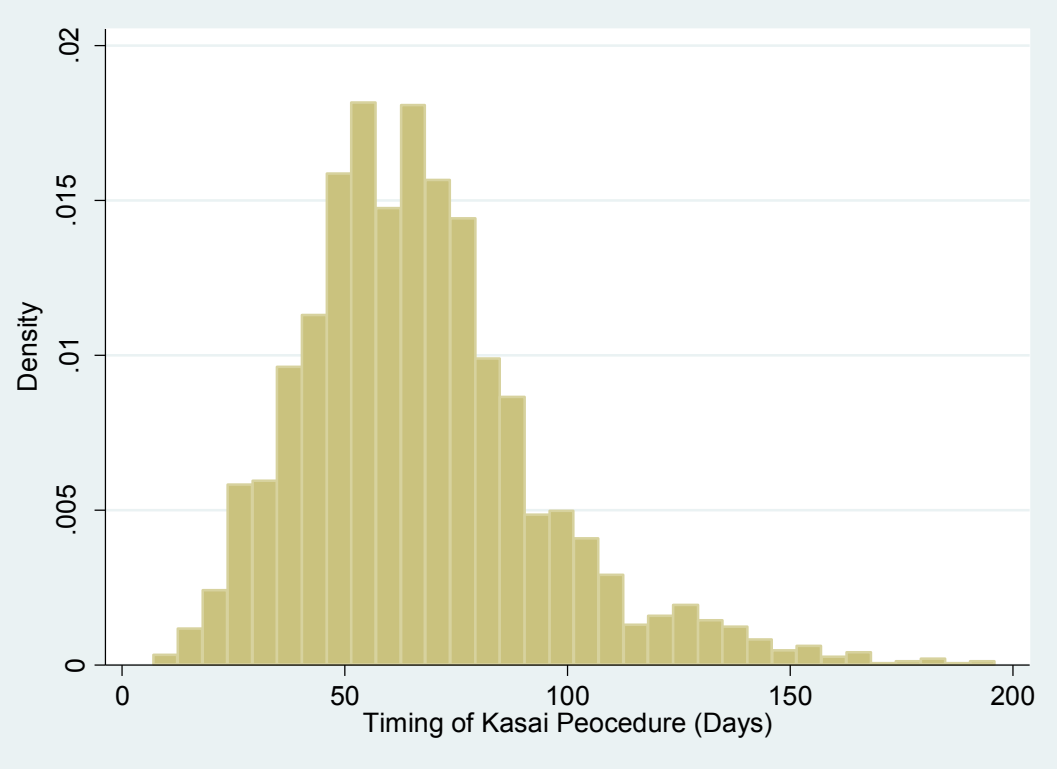


Figure 12 ROC with Full Covariates

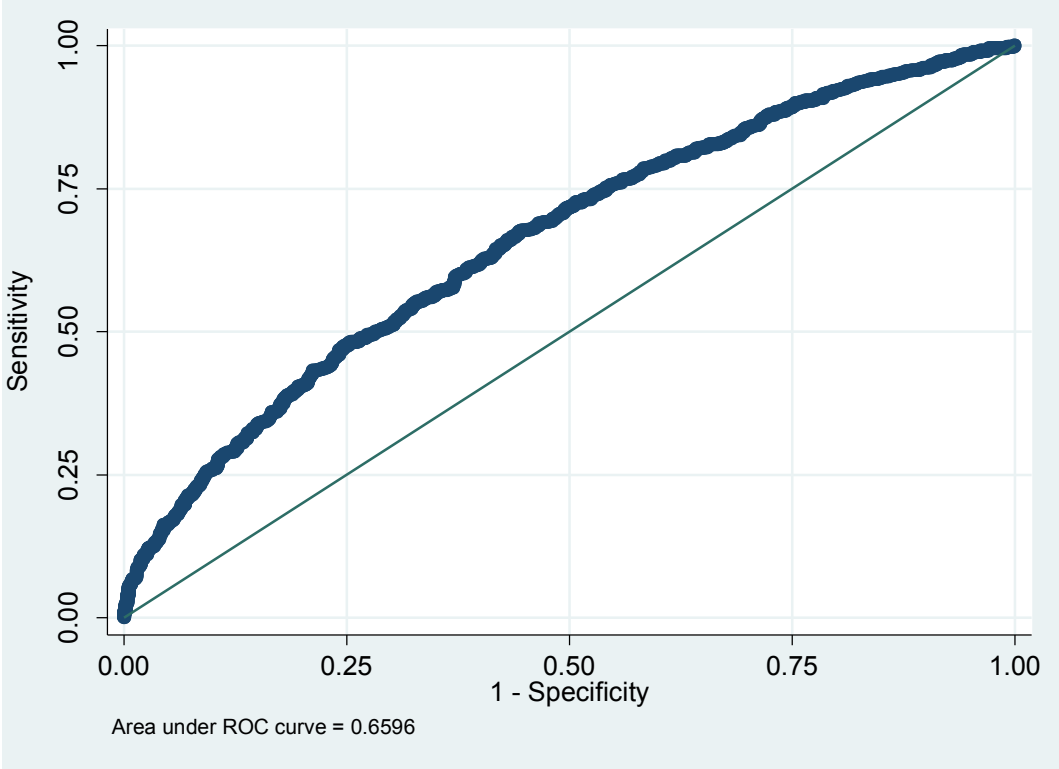


Figure 13 ROC with Full Covariates

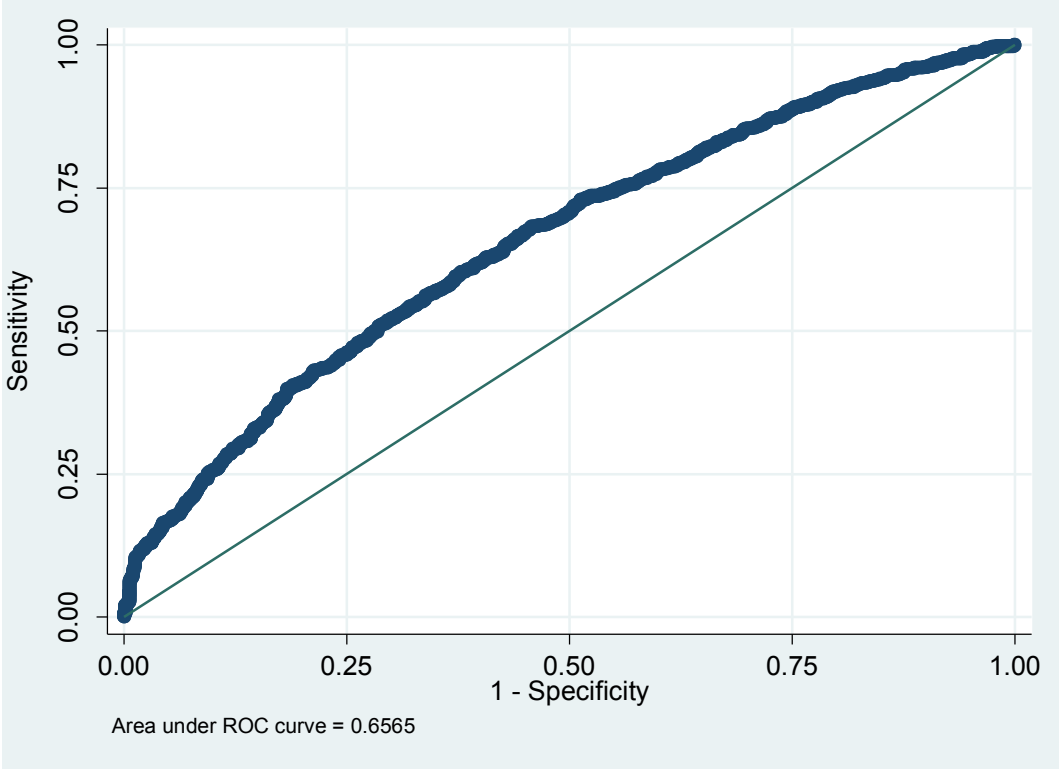


Figure 14 1-year Survival of Native Liver without Jaundice and Age at Kasai in Type I (n=91), Type I-cyst (n=199), Type II (n=47), and Type III (n=2065)

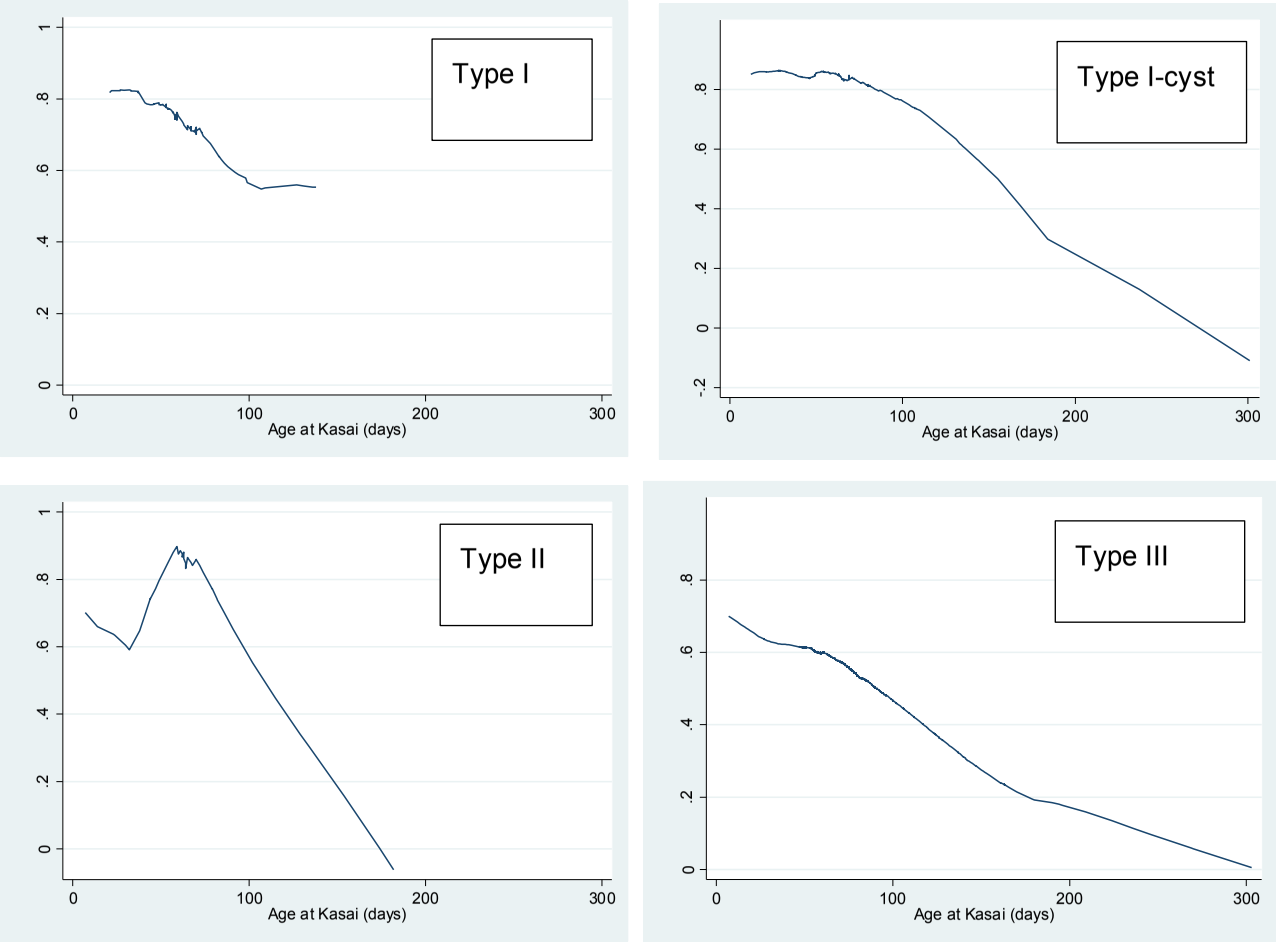


Figure 15 Plots of Regression Coefficients with Confidence Intervals

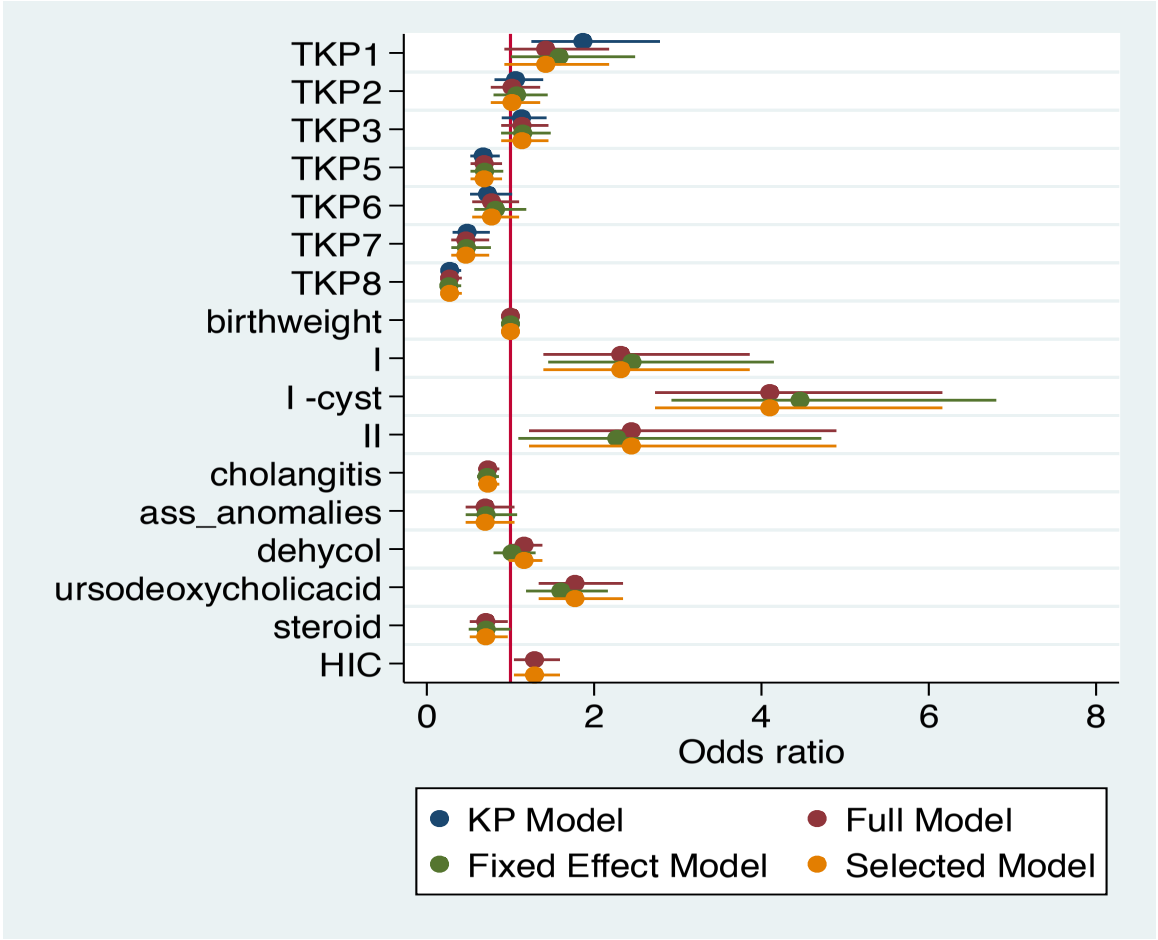


Table 1 Description of Associated Anomalies and Choleretic Agents

Terminologies	Description	Previously suggested effect on prognosis
Associated anomalies		
situs inversus	complete transposition (right to left reversal) of the thoracic and abdominal organs	having at least one of the left is classified as BA with Splenic Malformation and occurs 10-15% of infants with BA and show poorer outcomes (Dillon et al., 1994; Davenport, 2006)
asplenia	absence of a spleen or one that functions.	
polysplenia	a congenital disease manifested by multiple small accessory spleens	
accessory spleen	a small mass of splenic tissue that is separated from a patient's primary spleen	
preduodenal portal vein	a rare vascular anomaly that is the result of a variation in the normal developmental pattern of the embryonic precursors of the portal vein, the right and left vitelline veins and their three anastomotic channels	
Choleretic agents: leading to the stimulation of bile production and solids in the liver		
dehydrocholic	a synthetic bile acid, manufactured by the oxidation of cholic acid and acting as a hydrocholeretic, increasing bile output to clear increased bile acid load	not investigated
ursodeoxycholic acid	a hydrophilic bile acid; when given by mouth it shifts the balance of bile acids towards hydrophilic forms	positive (Meyers et al., 2003)
steroid (corticosteroid)	anti inflammatory, immunosuppressive, and/or choleretic effect but not known thoroughly	negative (Squires et al., 2014)

Table 2 Summary of Literature on Timing of Kasai Procedure and Survival of Native Liver

Authors	Year	Country	No. of cases	Age at Kasai (days)	Survival of Native Liver (%)				Statistical Test	*p*-value	Covariets (if any)
Tiao et al	1996-2003	Taiwan	327		5y				Kaplan–Meier, log–rank test	0.017	
				<60	54%						
				>60	37%						
Lien et al	1990-2003	Taiwan	191		3y	5y			logistic regression	<0.001	stool color card screening, prophylactic antibiotics, sex
				<60	57%	46.70%			covarites:		
				>60	29%	25.10%					
Hsiao et al	2004-2005	Taiwan	74		(jaudince free rate)				chi-square test, Fisher's exact test, Student t		
				<60	72%					0.002	
				>60	33%						
Schreiber et al	1985-2002	Canada	349		4y	10 y			rank-sum, log-rank, Kaplan-Meier	<0.0001	
				≤30days	49%	49%					
				31 to 90	36%	25%					
				>90 days	23%	15%					
Carvalho et al	1982-2008	Brazil	513		4y						
				<60	54%						
				61– 90	33%						
				>90	26%						
Wildhaber et al	1994-2004	Switzerland	48		4y				Kaplan-Meier	<0.001	
				<45	75%				log rank test		
				46-75	33%						
				>75	11%						
Chardot et al (1999)	1986 to 1996	France	472		5y	10 y			Kaplan-Meier	0.002	anatomical pattern,

				<=45	41%	41%			Cox regression		associated anomalies, caseload at hospital
				>45	31%	26%					
Chardot et al (2001)	1986 to 1996	France	472		5y	10 y			Kaplan-Meier	0.03	
				<=90	35%	30%			log-rank		
				>90	25%	22%					
Chardot et al (2013)	1986 to 2009	France	1107		5y	10 y	15y	20y	Kaplan-Meier	<0.001	anatomical pattern, associated anomalies
				<=30	54%	48%	39%	39%	log-rank		
				31-60	44%	40%	36%	32%	Cox regression		
				61-90	36%	32%	30%	28%			
				>90	31%	26%	19%	19%			
Serinet et al	1986-2002	France	743		2y	5y	10 y	15 y	Kaplan-Meier	<0.0001	
				<30	66%	58.10%	42.50%	-	log-rank		
				31–45	66%	40.50%	40.50%	40.50%			
				36-60	58%	42.40%	36.10%	33.30%			
				61-75	57%	36.10%	32.30%	28.70%			
				76-90	52%	26.40%	19.50%	16.70%			
				>90	42%	27.30%	20%	13.40%			
De Vries et al	1987-2008	Netherlands	104		20y				Kaplan-Meier	0.03 (60–75 d compared)	
				<45	14%				log rank test		
				45– 60	33%						
				60– 75	42%						
				>75	11%						
Karrer et al	1967 -1989	USA	904		Actual survival						
					5y				Kaplan-Meier		
				<30	63%				log-rank	0.023	
				31– 60	44%						
				61– 90	40%						
				91– 120	29%						
				>120	29%						

Table 3 Jaundice-Free Survival with Native Liver in Patients with BA in Taiwan Study

	3-Year Jaundice-Free* Survival with Native Liver	OR (95% CI)	*P* Value	5-Year Jaundice-Free Survival with Native Liver	OR (95% CI)	*P* Value
Age at Kasai operation			<0.001			0.02
≤60 days	63/111 (56.8%)	3.25 (1.76-6.00)		28/60 (46.7%)	2.63 (1.19-5.78)	
>60 days	23/80 (28.8%)	1.00		14/56 (25%)	1.00	
Stool card program			0.001			0.001
Cohort B+C†	58/102 (56.9%)	2.87 (1.58-5.21)		18/28 (64.3%)	4.80 (1.94-11.86)	
Cohort A‡	28/89 (31.5%)	1.00		24/88 (27.3%)	1.00	
Prophylactic antibiotics			0.009			0.01
Used	70/132 (53.0%)	3.03 (1.56-5.88)		30/65 (46.2%)	2.79 (1.25-6.20)	
No use	16/59 (27.1%)	1.00		12/51 (23.5%)	1.00	
Sex			0.44			0.58
Male	45/94 (47.9%)	1.25 (0.71-2.22)		21/54 (38.9%)	1.24 (0.58-2.65)	
Female	41/97 (42.3%)	1.00		21/62 (33.9%)	1.00	

*Total serum bilirubin <2.0 mg/dL.

†The data of cohort B and cohort C were merged to represent the outcomes in the era of the stool color card screening program. Cohort B (n = 28) includes patients born during the regional study of the stool color card screening program (2002-2003). Cohort C (n = 74) includes patients born after the launch of the nationwide study of the stool color card screening program (2004-2005).

‡Cohort A (n = 89) includes patients born before the institution of the stool color card screening program (1990-2000). Only 88 patients completed the 5-year follow-up.

Source: Lien et al (2011)

Table 4 Multivariate Analysis of Survival of Native Liver Outcome in France Study (1999)

Variable	RR	CI 95%	P
Survival with native liver			
Kasai operation			
Age ≤45 days	1	—	
Age >45 days	1.4	(1.03; 1.9)	.0002
Not performed	3.6	(2.0; 6.2)	
Anatomical pattern of atresia			
(1) Atresia limited to common bile duct			
(2) Atresia with cyst in the liver hilum and hairy intrahepatic ducts			
(3) Atresia with gallbladder and common bile duct patent			
(4) Complete extrahepatic biliary atresia			
Patterns (1) and (2)	1	—	<.0001
Patterns (3) and (4)	4.1	(2.1; 8.0)	
Polysplenia syndrome			
No	1	—	
Yes	1.7	(1.1; 2.5)	.01
Experience of the center (mean 1986-1996)			
≥20 Kasai operations/year	1	—	.0001
≤5 Kasai operations/year	1.6	(1.3; 2.1)	

Source: Chardot et al (1999)

Table 5 Univariate and Multivariate Analysis of Survival of Native Liver in France Study (2013)

A

Prognostic factor	N patients	Survival with native liver (SE: standard error) [number of patients alive with native liver at follow-up]				*p* values
		5-year SNL	10-year SNL	15-year SNL	20-year SNL	
Anatomical pattern of the extrahepatic biliary remnant						<0.0001
Type 1	24	87.3% (SE: 6.9%) [18]	82.5% (SE: 8.0%) [11]	82.5% (SE: 8.0%) [6]	82.5% (SE: 8.0%) [2]	
Type 2	75	60.9% (SE: 6.3%) [26]	55.4% (SE: 6.8%) [18]	51.9% (SE: 7.2%) [11]	51.9% (SE: 7.2%) [5]	
Type 3	173	47.1% (SE: 4.0%) [67]	42.5% (SE: 4.0%) [39]	37.2% (SE: 4.3%) [19]	37.2% (SE: 4.3%) [8]	
Type 4	719	34.9% (SE: 1.9%) [199]	31.0% (SE: 1.9%) [108]	27.2% (SE: 2%) [52]	23.5% (SE: 2.5%) [15]	
BA splenic malformation syndrome						<0.0001
Absent	829	42.8% (SE: 1.8%) [287]	38.6% (SE: 1.8%) [167]	34.2% (SE: 1.9%) [82]	31.2% (SE: 2.3%) [28]	
Present	86	19.6% (SE: 4.7%) [12]	15.1% (SE: 4.6%) [6]	15.1% (SE: 4.6%) [4]	15.1% (SE: 4.6%) [2]	
Age at Kasai operation						0.0002
≤30 days	99	53.5% (SE: 5.2%) [37]	48.2% (SE: 5.6%) [20]	38.9% (SE: 7.5%) [6]	38.9% (SE: 7.5%) [1]	
31 to 60 days	435	43.5% (SE: 2.5%) [152]	39.6% (SE: 2.5%) [86]	35.9% (SE: 2.7%) [48]	31.7% (SE: 3.4%) [14]	
61 to 90 days	361	35.5% (SE: 2.6%) [99]	31.8% (SE: 2.7%) [58]	29.8% (SE: 2.7%) [30]	28.1% (SE: 3.1%) [15]	
>90 days	132	31.0% (SE: 4.3%) [30]	26.1% (SE: 4.3%) [14]	18.7% (SE: 4.8%) [6]	18.7% (SE: 4.8%) [1]	

B

Prognostic factor	Relative risk	95% CI	*p* values
Anatomical pattern of the extrahepatic biliary remnant			<0.0001
Type 1	0.128	0.041-0.399	
Type 2	0.474	0.319-0.705	
Type 3	0.685	0.539-0.870	
Type 4	1		
BA splenic malformation syndrome			0.0001
Absent	0.593	0.453-0.776	
Present	1		
Age at Kasai operation			<0.0001
1st month of life (age <31 days)	0.540	0.372-0.785	
2nd month of life (age 31 to 60 days)	0.583	0.452-0.752	
3rd month of life (age 61 to 90 days)	0.744	0.372-0.785	
4th month of life and after (age >90 days)	1		

Source: Chardot et al (2013)

Table 6 Summary of Literature on Timing of Kasai Procedure and Survival of Native Liver

Author(s) and Year Published	Country	Data	Timing of Height for Age	Educational Outcome	Endogeneity	Results
Victora et al. (2008)	Brazil, Guatemala, India, the Philippines, and South Africa	Brazil (cohort 1982), Guatemala (Community Trial 1969-77), India (Cohort 1969-72), the Philippines (Cohort 1983-84), and South Africa(Cohort 1990)	At roughly 2	Years of schooling	Ordinary Least Square (OLS) with controlling for age, parents' educational level, and child's socioeconomic status	One unit increase in height-for-age (HA), 0.5 increased years of schooling ($p<0.01$)
Glewwe and Jacoby (1995)	Ghana	Cross-sectional data, 1757 Ghanaian children aged 6–15 years (1989)	6-15	Primary school entrance	Family Fixed Effect (FFE), Instrumental Variable (IV) (distance to nearby hospitals and maternal height)	With a decline in HA of 1, 0.634 month decrease in enrollment and the t-value is 2.49. Insignificant effect on school attainment as measured by grade completed.
Glewwe, Jacoby, and King (2001)	The Philippines	Panel data, 2192 households (1984)	8	Test score	FFE+ IV (the height of the older child at age 24 months)	One unit increase in HA raises the achievement test score by 8.9 (2.8) points. Such effect is equivalent to spending about eight extra months in school
Alderman, Behrman, Lavy, and Menon(2001)	Pakistan	Panel data, 800 households (1986 to 1991)	5	Probability of being enrolled in school at age 7	IV (deviations in prices from long-term trends)	One unit increase in HA increases the probability of being enrolled in school at age 7, especially for girls (0.51 with SE=0.14 for girls; 0.07 with SE=0.18 for boys). These results are equivalent to an improvement of 0.25 in HA score raise the probability of school enrollment for girls by 9 percent over the base case and for boys by 2 percent.
Alderman, Hoddinott, and Kinsey (2006)	Zimbabwe	Panel data, 665 young adults (1983/1984 and 1987/2000)	5	1) Years of schooling 2) Delayed enrollment status	FFE (siblings)+ IV (representations of civil war and drought Civil war)	One unit increase in HA increases number of grades of schooling completed by 0.678 (t=2.13) as well as starting school at a younger age

					status	(0.400 (t=1.65)). It also leads to increased height as a young adult. Exposure to civil war and drought decreases in child HA by 0.049 (t=3.17).
INCAP study (Behrman and others 2009; Hoddinott and others 2008; Maluccio and others 2009)	Guatemala	1969 to 1977	36 months	1) Grade attainment 2) Reading comprehension	Shock measure as randomly assigned high-protein energy drink (atole) and placebo (fresco) in 1969–1977 Difference-in-difference (DiD), Village Fixed Effect, 2SLS with instrument as an exposure to intervention	Female schooling increased by 1.17 grade (t=2.13) Reading comprehension increased by 0.28 scores (t=2.52)
Alderman, Hoogeveen, and Rossi (2009)	Tanzania	1991-1994, 2004	As a percentage of the median of the reference population from 0-10	1) Delayed enrollment 2) Years of schooling	Community fixed effect, IV as crop loss and flood	One unit increase in HA decrease 0.591(z=3.42) delayed enrollment and one unit increase in HA increase the 0.575 years of schooling (z=2.18). Similar results attained for OLS.
Linnemayr & Alderman (2011)	Senegal	Data from the program's two survey rounds (2004 and 2006), (10,127 obs.)	*Weight for age is measured instead at age 5		Nutrition enhancement program: vitamin A and deworming for children 6–59 months, iron for pregnant women, bednets, breastfeeding promotion, cooking workshops, 2004–2006 OLS, 2SLS (instrument as planned treatment status), a combination of DiD and Propensity Score Matching (PSM)	Insignificant impact of planned treatment status on z scores. DiD and PSM estimates show 0.27 standard deviation increase in weight-for-age score (p=0.05).

Table 7 Calculation of DiD

	Exposed group	Control Group	Difference across groups
T1	$Y_{t=1}$	$C_{t=1}$	$Y_{t=1} - C_{t=1}$
T0	$Y_{t=0}$	$C_{t=0}$	$Y_{t=0} - C_{t=0}$
Difference Across time	D1: $Y_{t=1} - Y_{t=0}$	D2: $C_{t=1} - C_{t=0}$	DiD: $(Y_{t=1} - Y_{t=0}) - (C_{t=1} - C_{t=0})$

Table 8 Estimated Number of BA Patients in Japan

Years	Number of Birth (1,000)	Number of Birth	Number of Patients when Incidence (1/9640)
1989	1,247	1247000	108
1990	1,222	1222000	108
1991	1,223	1223000	108
1992	1,209	1209000	108
1993	1,188	1188000	108
1994	1,238	1238000	108
1995	1,187	1187000	108
1996	1,207	1207000	108
1997	1,192	1192000	108
1998	1,203	1203000	108
1999	1,178	1178000	108
2000	1,191	1191000	108
2001	1,171	1171000	108
2002	1,154	1154000	108
2003	1,124	1124000	108
2004	1,111	1111000	108
2005	1,063	1063000	108
2006	1,093	1093000	108
2007	1,090	1090000	108
2008	1,091	1091000	108
2009	1,070	1070000	108
2010	1,071	1071000	108
2011	1,051	1051000	108
2012	1,037	1037000	108
		Estimated Number of Patients	2864

Calculations were based on birth statistics from Ministry of Health, Labor and Welfare and the rate of incidence of BA reported by Nio M et al (2003).

Table 9 Baseline Characteristics of Patients

Variables	n(%)
Sociodemographics	
Sex	
Male	956 (37)
Female	1643 (63)
Birth order, median (IQR)	2 (1-2)
Birth weight (g) , mean (±SD)	2916.9 (449)
Gestational age (wks) , mean (±SD)	38.7 (1.7)
Age of Parents	
Age of Father (yrs), mean (±SD)	31.9 (5.8)
Age of Mother (yrs), mean (±SD)	29.5 (4.6)
Clinical Characteristics	
Age at Kasai (days), mean (±SD)	67.6 (29.3)
Types of obstruction[a]	
I	101 (4)
I-cyst	212 (8)
II	54 (2)
III	2215 (85)
Associated anomalies	129 (5)
Cholangitis	1046 (42)
Use of choleretic agent	
Dehydrocholic	1321 (50)
Ursodeoxycholic Acid	2270 (87)
Secretin	115 (4)
Steroid	2314 (88)
Glucagon	593 (22)
PGE1	131 (5)
Taurine	264 (10)
PG-F2α	49 (2)
PG-E2	229 (9)
Hospital Characteristics	
Caseload, mean (±SD)	50 (31)
1-year Survival With Native Liver without Jaundice	1551 (60)

Table 10 Screening Method used Prior to Kasai Procedure

Screening Method	n(%)
Measurement of total bile acid	8(35)
Stool color card	4(16)
Lipoprotein X detection	3(15)
Others	8(34)
Total	23(100)

Table 11 Bivariate Analysis on Timing of Kasai Procedure by 1-Year Native Liver Survival without Jaundice

Timing of KP	1-year Survival of Native Liver without Jaundice		
	Yes	No	*p*-value
	n (%)	n (%)	
Less than 30	117 (7.6)	41 (3.9)	<0.001
31 to 45	231 (14.9)	135 (13)	0.163
46 to 60	407 (26.2)	213 (20.4)	0.001
61 to 75	398 (25.6)	241 (22.6)	0.082
76 to90	212 (13.7)	186 (17.7)	0.004
91 to 105	97 (6.3)	83 (7.8)	0.135
106 to 120	41 (2.6)	52 (4.9)	0.002
over 121	48 (3.1)	103 (9.7)	<0.001
Total	1551 (100)	1054 (100)	

Table 12 Bivariate Analysis of Socio demographics, Clinical and Hospital Characteristics by 1-Year Survival of Native Liver without Jaundice

Variables	1-Year Survival of Native Liver without Jaundice		
	Yes	No	*p*-value
Sociodemographics			
Gender (%)			
Male	580 (60.7)	376 (39.3)	0.38
Female	968 (58.9)	675 (41.1)	
Birth order, median	2	2	0.87
Birth weight (g) , mean	2927	2902	0.15
Gestational age (wks) , mean	38.7	38.7	0.66
Age of Parents			
Age of Father (yrs), mean	31.9	32.1	0.31
Age of Mother (yrs), mean	29.5	29.5	0.87
Clinical Characteristics			
Age at Kasai (days), mean	63	74	<0.000
Type of obstruction			<0.001
I (%)	74 (73.3)	27 (26.7)	
I-cyst (%)	175 (82.6)	37 (17.5)	
II (%)	40 (74.1)	17 (25.9)	
III (%)	1251 (56.5)	964 (45.5)	
Associated anomalies (%)	65 (4)	64 (6)	0.03
Cholangitis (%)	593 (39)	453 (47)	<0.001
Use of choleretic agent			
Dehydrocholic (%)	819 (53)	502 (47)	0.006
Ursodeoxycholic Acid (%)	1394 (90)	876 (82)	<0.001
Secretin (%)	63 (4)	52 (5)	0.30
Steroid (%)	1382 (89)	932 (88)	0.36
PGE1 (%)	77 (5)	54 (5)	0.89
Taurine (%)	164 (11)	100 (9)	0.34
PG-F2α (%)	26 (2)	23 (2)	0.36
PG-E2 (%)	126 (8)	103 (10)	0.16
Hospital Characteristics			
High Caseload[a], (%)	351 (23)	196 (18)	0.01

Table 13 Multivariate Logit and Multivariate Probit Regression on Timing of Kasai Procedure by 1-year Survival of native Liver without Jaundice

Variables			
	Logit	**Probit**	**Ratio**
Age at Kasai Procedure (days)			
Less than 30	0.553**(2.78)	0.335**(2.83)	1.65
31 - 45	0.0153 (0.11)	0.00947 (0.11)	1.62
46 - 60	0.127 (1.09)	0.0784 (1.09)	1.62
61 - 75 (reference)	-	-	
76 - 90	-0.385**(-2.98)	-0.240**(-2.98)	1.60
91- 105	-0.338*(-1.99)	-0.210*(-1.98)	1.61
106 - 120	-0.742***(-3.31)	-0.463***(-3.31)	1.60
Over 121	-1.299***(-6.40)	-0.807***(-6.53)	1.61
Constant	0.504***(-6.18)	0.315***(-6.24)	1.60
N	2605	2605	

t statistics in parentheses ="* p<0.05 ** p<0.01 *** p<0.001"

Table 14 Covariates Selection

Variables	Included in multivariate analysis or not	Previously suggested as confounding	*p*-value in bivariate analysis	10% change in the main effect (Base OR on KP .82)
Sociodemographics				
Gender	×	×	0.38	×
Birth order	×	×	0.87	×
Birth weight	○	×	0.15	-
Gestational age	×	×	0.66	×
Age of Parents				
Age of Father	×	N/A	0.31	×
Age of Mother	×	N/A	0.87	×
Clinical Characteristics				
Type of obstruction				
I	○	○	0.004	-
I-cyst	○	○	<0.001	-
II	○	○	0.03	-
III	○	○	<0.001	-
Associated anomalies	○	○	0.03	-
Cholangitis	○	○	<0.001	-
Use of choleretic agent				
Dehydrocholic	○	N/A	0.006	-
Ursodeoxycholic Acid	○	○	<0.001	-
Secretin	×	×	0.30	×
Steroid	○	○	0.36	×
PGE1	×	×	0.89	×
Taurine	×	×	0.34	×
PG-F2α	×	×	0.36	×
PG-E2	×	×	0.16	×
Hospital Characteristics				
High Caseload	○	○	0.01	-

Table 15 Multivariate Logit Regression on Timing of Kasai Procedure by 1-year Survival of native Liver without Jaundice with Full Covariates vs Selected Covariates

Variables				
	No Covariates	Full Covariates	Selected Covariates	Selected Covariates + Hospital FE
Age at Kasai Procedure (days)				
Less than 30	0.553**(2.78)	0.2 (0.87)	0.322 (1.49)	0.458* (1.98)
31 - 45	0.0153 (0.11)	0.00328 (0.02)	0.00641 (0.04)	0.0706 (0.47)
46 - 60	0.127 (1.09)	0.072 (0.55)	0.134 (1.07)	0.137 (1.05)
61 - 75 (reference)	-	-	-	
76 - 90	-0.385**(-2.98)	-0.418** (-2.84)	-0.376** (-2.73)	-0.370** (-2.59)
91- 105	-0.338*(-1.99)	-0.338 (-1.74)	--0.26 (-1.43)	-0.197 (-1.04)
106 - 120	-0.742***(-3.31)	-0.836*** (-3.41)	-0.767** (-3.21)	-0.747** (-3.05)
Over 121	-1.299***(-6.40)	-1.312*** (-5.66)	-1.330*** (-6.07)	-1.335*** (-5.89)
Gender	-	0.133 (1.38)	-	-
Birth order	-	0.0118 (0.21)	-	-
Birth weight (g)	-	0.00003 (0.26)	0.00006 (0.70)	-0.00009 (-0.96)
Gestational age (wks)	-	-0.0222 (-0.68)	-	-
Age of Parents				
Age of Father	-	-0.00271(-0.24)	-	-
Age of Mother	-	-0.0157(-1.15)	-	-
Clinical Characteristics				
Type of obstruction				
I	-	0.757** (2.87)	0.842**(3.24)	0.897*** (3.35)
I-cyst	-	1.487*** (6.59)	1.420***(6.83)	1.495*** (6.94)
II	-	0.763* (2.1)	0.897*(2.53)	0.820* (2.2)
III (reference)		-	-	-
Associated anomalies	-	-0.411 (-1.95)	-0.384 (-1.86)	-0.345 (-1.60)
Cholangitis	-	-0.254** (-2.74)	-0.307***	-0.329*** (-3.56)

			(-3.50)	
Use of choleretic agent				-
Dehydrocholic	-	0.187* (1.99)	0.155 (1.76)	0.0181 (0.14)
Ursodeoxycholic Acid	-	0.587*** (3.9)	0.566*** (3.97)	0.471** (3.07)
Secretin	-	-0.163 (-0.74)	-	-
Steroid	-	-0.413* (-2.40)	-0.340* (-2.11)	-0.344 (-1.94)
PGE1	-	-0.0525 (-0.25)	-	-
Taurine	-	0.0371 (0.24)	-	-
PG-F2α	-	-0.398 (-1.23)	-	-
PG-E2	-	-0.188 (-1.17)	-	-
Hospital Characteristics				
High Caseload		0.219 (1.89)	0.255* (2.36)	-
Constant	0.504***(-6.18)	1.79 (1.49)	0.482 (1.42)	-
N	2605	2182	2422	2377
Pseudo R2	0.025	0.059	0.058	-

t statistics in parentheses ="* p<0.05 ** p<0.01 *** p<0.001"

Table 16 Multivariate Logit Regression on Timing of Kasai Procedure by 1-year Survival of Native Liver without Jaundice with Selected Covariates

Variables	**Type I (CI)**	**Type I-cyst (CI)**	**Type II (CI)**	**Type III (CI)**
Age at Kasai Procedure (days)				
Less than 30	-0.6 (-3.4 - 2.2)	-0.0981 (-1.8 - 1.6)	-1.678 (-4.6 - 1.2)	0.43 (-.0 - .9)
31 - 45	1.256 (-.7- 3.3)	-0.115 (-1.8 - 1.6)	-1.819 (-4.9 - 1.3)	0.0276 (-.3 - .3)
46 - 60	1.02 (-.9- 2.9)	-0.84 (-2.4 - .7)	0.0919 (-3.2 - 3.4)	0.145 (-.1- .4)
61 - 75 (reference)	-			-
76 - 90	-1.176 (-3.4 - 1.0)	0.225 (-1.8 - 2.3)	-3.836* (-7.5 -.1)	-0.359* (-.6 -.1)
91- 105	-1.628 (-4.9 - 1.6)	-0.962 (-2.9 - 1.0)	-	-0.213 (-.6 - .2)
106 - 120	0.266 (-2.8 - 3.3)	-0.611 (-3.2 – 2.0)	-	-0.832** (-1.3 -.3)
Over 121	-0.209 (-3.5 - 3.1)	-2.459** (-4.2 - .7)	-4.645* (-8.5 -.8)	-1.175*** (-1.6 -.7)
Sociodemographics				
Birth weight (g)	-0.000575 (-.001 - .001)	-0.00004 (-.0 - .0)	-0.00005 (-.0 - .0)	-0.00002 (-.0 - .0)
Clinical Characteristics				
Associated anomalies	0.406 (-1.8 - 2.6)	-1.766 (-3.5 - .0)	-	-0.337 (-.8 -.1)
Cholangitis	-1.967* (-3.5 - -.4)	-1.476** (-2.4 - -.5)	-1.049 (-3.2 - 1.1)	-0.232* (-.4 -.1)
Use of choleretic agent				

Use of choleretic agent				
Dehydrocholic	0.464 (-.8 - 1.7)	0.353 (.6 - 1.3)	-1.169 (-3.6 - 1.2)	0.146 (-.0 -.3)
Ursodeoxycholic Acid	4.274** (1.2 - 7.3)	1.611* (.6 - 1.3)	-1.521 (-4.9 - 1.9)	0.541*** (.2 - .8)
Steroid	-4.858* (-8.8 - -.9)	-1.539 (-3.3 - .3)	-0.0662 (-3.7 - 3.6)	-0.232 (-.6 - .1)
Hospital Characteristics				
High Caseload	0.464 (-1.2 - 2.2)	0.713 (-.6 - 2.0)	1.558 (-1.3 - 4.4)	0.253* (.0 - .5)
Constant	3.959 (-1.6 - 9.5)	2.951 (-1.69- 4.1)	6.439 (-1.7 - 14.6)	0.24 (-.5 - .9)
N	91	199	47	2065
Pseudo R2	0.25	0.18	0.36	0.36

t statistics in parentheses ="* p<0.05 ** p<0.01 *** p<0.001"

Table 17 1st Stage OLS and Probit Estimates of 1-Year Native Liver Survival without Jaundice of Type III Adjusted for Sociodemographics and Clinical and Hospital Characteristics

Age at Kasai Procedure	<30	<45	< 60	<75	<90	<105	<120
OLS							
Shorter Length of Hospital Stay	-0.0341** (-3.27)	-0.114*** (-6.53)	-0.185*** (-8.62)	-0.0805*** (-3.98)	-0.299 (-1.87)	-0.0177 (-1.44)	-0.0076 (-0.81)
Probit							
Shorter Length of Hospital Stay	-0.0341** (-3.27)	-0.114*** (-6.54)	-0.185*** (-8.65)	-0.0805*** (-3.99)	-0.0299 (-1.88)	-0.0177 (-1.45)	-0.0077 (-0.81)
F statistics	10.668	42.589	74.382	15.833	3.514	2.087	0.659
N	2405	2405	2405	2405	2405	2405	2405

t statistics in parentheses ="* p<0.05 ** p<0.01 *** p<0.001"

Table 18 Multivariate Probit Regression on Timing of Kasai Procedure by 1-Year Native Liver Survival without Jaundice with OLS vs Probit +OLS IV vs Probit+ IV

Variables						
	OLS	**Probit**	**OLS**	**Probit**	**OLS+ IV**	**Probit+IV**
Age at Kasai Procedure (days)						
<30	0.160***(3.99)	0.446***(4.01)	0.0968*(2.36)	0.295* (2.46)	**0.380 (0.61)**	**1.07 (0.67)**
<45	0.0858***(3.58)	0.226***(3.59)	0.0611*(2.50)	0.174* (2.57)	**0.114 (0.62)**	**0.333 (0.67)**
<60	0.110***(5.67)	0.285*** (5.65)	0.0944***(4.81)	0.258*** (4.8)	**0.0699 (0.62)**	**0.198 (0.64)**
<75	0.160***(7.74)	0.409***(7.61)	0.140***(6.69)	0.377*** (6.61)	**0.161 (0.62)**	**0.453 (0.64)**
<90	0.180*** (6.86)	0.457*** (6.71)	0.158***(5.89)	0.421*** (5.83)	0.433 (0.61)	1.166 (0.7)
<105	0.251*** (7.43)	0.640*** (7.19)	0.238***(6.88)	0.635*** (6.71)	0.729 (0.60)	1.908 (0.75)
<120	0.296*** (6.83)	0.763*** (6.57)	0.283***(6.28)	0.762*** (6.13)	1.681 (0.52)	3.604 (1.06)
Sociodemographics			✓	✓	✓	✓
Clinical Characteristics			✓	✓	✓	✓
Hospital Characteristics			✓	✓	✓	✓
N	2605	2605	2405	2405	2405	2405

t statistics in parentheses ="* p<0.05 ** p<0.01 *** p<0.001"

Table 19 OLS Result of Timing of Kasai Procedure on 1-Yesr Native Liver Survival

	KP<30	KP<45	KP<60	KP<75	KP<90	KP<105	KP<120
TKP30	0.0968*						
	(2.36)						
TKP45		0.0611*					
		(2.50)					
TKP60			0.0944***				
			(4.81)				
TKP75				0.140***			
				(6.69)			
TKP90					0.158***		
					(5.89)		
TKP105						0.238***	
						(6.88)	
TKP120							0.283***
							(6.28)
birthweight	0.000000375	0.000000335	-0.00000292	-0.00000790	-0.00000568	-0.00000605	-0.00000588
	(0.02)	(0.02)	(-0.13)	(-0.36)	(-0.26)	(-0.28)	(-0.27)
i	0.185***	0.183***	0.171***	0.173***	0.178***	0.183***	0.183***
	(3.63)	(3.59)	(3.35)	(3.41)	(3.50)	(3.60)	(3.60)
icyst	0.248***	0.252***	0.252***	0.261***	0.265***	0.270***	0.268***
	(6.87)	(7.02)	(7.08)	(7.40)	(7.48)	(7.65)	(7.58)
ii	0.185**	0.188**	0.189**	0.174*	0.182**	0.185**	0.194**
	(2.72)	(2.76)	(2.78)	(2.58)	(2.68)	(2.74)	(2.87)
cholangitis	-0.0720***	-0.0720***	-0.0720***	-0.0719***	-0.0720***	-0.0701***	-0.0728***
	(-3.65)	(-3.65)	(-3.67)	(-3.68)	(-3.67)	(-3.59)	(-3.72)
ass_anomalies	-0.0686	-0.0720	-0.0758	-0.0875	-0.0744	-0.0691	-0.0691
	(-1.46)	(-1.53)	(-1.62)	(-1.87)	(-1.59)	(-1.48)	(-1.48)
HIC	0.0558*	0.0558*	0.0593*	0.0527*	0.0517*	0.0541*	0.0544*
	(2.34)	(2.35)	(2.50)	(2.24)	(2.19)	(2.29)	(2.30)
dehycol	0.0351	0.0364	0.0337	0.0330	0.0327	0.0354	0.0359
	(1.79)	(1.85)	(1.72)	(1.69)	(1.67)	(1.82)	(1.84)
ursodeoxycholic acid	0.133***	0.132***	0.127***	0.129***	0.129***	0.130***	0.129***
	(4.13)	(4.10)	(3.97)	(4.03)	(4.02)	(4.08)	(4.03)
steroid	-0.0759*	-0.0727*	-0.0697*	-0.0641	-0.0735*	-0.0749*	-0.0804*
	(-2.19)	(-2.10)	(-2.02)	(-1.87)	(-2.14)	(-2.19)	(-2.34)
_cons	0.532***	0.523***	0.505***	0.460***	0.425***	0.337***	0.293***
	(7.25)	(7.13)	(6.90)	(6.27)	(5.68)	(4.34)	(3.59)
N	2405	2405	2405	2405	2405	2405	2405
t statistics in parentheses							
=″* p<0.05	** p<0.01	*** p<0.001″					

Table 20 Probit Result of Timing of Kasai Procedure on 1-Yesr Native Liver Survival

	KP<30	KP<45	KP<60	KP<75	KP<90	KP<105	KP<120
	0.295*						
KP<30	−2.46						
		0.174*					
KP<45		−2.57					
			0.258***				
KP<60			−4.8				
				0.377***			
KP<75				−6.61			
					0.421***		
KP<90					−5.83		
						0.635***	
KP<105						−6.71	
							0.762***
KP<120							−6.13
	1.3E−06	3.83E−07	−8E−06	−2.2E−05	−1.6E−05	−1.6E−05	−1.6E−05
birthweight	−0.02	−0.01	(−0.13)	(−0.36)	(−0.27)	(−0.27)	(−0.27)
	0.541***	0.535***	0.508***	0.515***	0.525***	0.539***	0.537***
i	−3.6	−3.56	−3.36	−3.4	−3.49	−3.58	−3.56
	0.787***	0.799***	0.800***	0.834***	0.847***	0.865***	0.862***
icyst	−6.89	−7.02	−7.05	−7.32	−7.41	−7.54	−7.49
	0.535**	0.544**	0.550**	0.525*	0.535**	0.551**	0.579**
ii	−2.67	−2.71	−2.73	−2.57	−2.65	−2.71	−2.84
	−0.200***	−0.199***	−0.201***	−0.203***	−0.202***	−0.197***	−0.205***
cholangitis	(−3.74)	(−3.73)	(−3.77)	(−3.78)	(−3.77)	(−3.68)	(−3.82)
	−0.188	−0.197	−0.208	−0.241	−0.206	−0.189	−0.19
ass_anomalies	(−1.49)	(−1.56)	(−1.65)	(−1.91)	(−1.63)	(−1.50)	(−1.51)
	0.153*	0.154*	0.164*	0.147*	0.144*	0.151*	0.152*
HIC	−2.35	−2.36	−2.51	−2.24	−2.2	−2.3	−2.32
	0.0976	0.102	0.0937	0.0916	0.091	0.0992	0.101
dehycol	−1.83	−1.9	−1.75	−1.71	−1.7	−1.85	−1.87
	0.358***	0.356***	0.345***	0.352***	0.351***	0.356***	0.352***
ursodeoxycholic acid	−4.11	−4.09	−3.96	−4.02	−4	−4.06	−4.02
	−0.219*	−0.212*	−0.204*	−0.191*	−0.217*	−0.222*	−0.238*
steroid	(−2.27)	(−2.19)	(−2.10)	(−1.97)	(−2.23)	(−2.29)	(−2.44)
	0.0855	0.0628	0.016	−0.101	−0.192	−0.427*	−0.554*
Intercept	−0.43	−0.32	−0.08	(−0.50)	(−0.94)	(−1.99)	(−2.44)
n	2405	2405	2405	2405	2405	2405	2405
t statistics in pa	** p<0.01	*** p<0.001"					
="* p<0.05							

Table 21 OLS +IV Result of Timing of Kasai Procedure on 1-Yesr Native Liver Survival

	KP<30	KP<45	KP<60	KP<75	KP<90	KP<105	KP<120
TKP30	0.380						
	(0.61)						
TKP45		0.114					
		(0.62)					
TKP60			0.0699				
			(0.62)				
TKP75				0.161			
				(0.62)			
TKP90					0.433		
					(0.61)		
TKP105						0.729	
						(0.60)	
TKP120							1.681
							(0.52)
birthweight	-0.00000760	-0.00000205	-0.00000135	-0.00000950	-0.0000210	-0.0000249	-0.0000503
	(-0.27)	(-0.09)	(-0.06)	(-0.32)	(-0.46)	(-0.48)	(-0.48)
i	0.179***	0.180***	0.175**	0.171**	0.161*	0.172**	0.158
	(3.36)	(3.42)	(3.22)	(2.98)	(2.36)	(2.94)	(1.90)
icyst	0.205*	0.242***	0.255***	0.261***	0.267***	0.285***	0.293***
	(2.02)	(4.90)	(6.70)	(7.36)	(7.28)	(5.53)	(4.17)
ii	0.171*	0.187**	0.189**	0.172*	0.167*	0.175*	0.214*
	(2.27)	(2.73)	(2.79)	(2.34)	(2.10)	(2.35)	(2.33)
cholangitis	-0.0664**	-0.0702***	-0.0725***	-0.0716***	-0.0684**	-0.0622*	-0.0673*
	(-2.84)	(-3.41)	(-3.68)	(-3.61)	(-3.12)	(-2.20)	(-2.54)
ass_anomalie	-0.0692	-0.0751	-0.0739	-0.0903	-0.0849	-0.0706	-0.0727
	(-1.46)	(-1.56)	(-1.55)	(-1.55)	(-1.55)	(-1.45)	(-1.30)
HIC	0.0583*	0.0566*	0.0582*	0.0524*	0.0461	0.0524*	0.0521
	(2.37)	(2.37)	(2.40)	(2.20)	(1.64)	(2.11)	(1.83)
dehycol	0.0325	0.0367	0.0343	0.0325	0.0268	0.0341	0.0353
	(1.58)	(1.87)	(1.74)	(1.61)	(1.07)	(1.66)	(1.53)
ursodeoxych	0.134***	0.131***	0.129***	0.128***	0.122***	0.125***	0.110*
	(4.12)	(4.08)	(3.96)	(3.93)	(3.33)	(3.58)	(1.97)
steroid	-0.0855*	-0.0728*	-0.0705*	-0.0629	-0.0752*	-0.0798*	-0.119
	(-2.10)	(-2.11)	(-2.04)	(-1.68)	(-2.13)	(-2.12)	(-1.22)
_cons	0.547***	0.519***	0.510***	0.451**	0.249	-0.0509	-0.859
	(6.72)	(7.01)	(6.61)	(3.19)	(0.54)	(-0.05)	(-0.33)
N	2405	2405	2405	2405	2405	2405	2405
t statistics in parentheses							
=″* p<0.05	** p<0.01	*** p<0.001″					

Table 22 Probit+IV Result of Timing of Kasai Procedure on 1-Yesr Native Liver Survival

	KP<30	KP<45	KP<60	KP<75	KP<90	KP<105	KP<120
TKP30	1.07						
	−0.67						
TKP45		0.333					
		−0.67					
TKP60			0.198				
			−0.64				
TKP75				0.453			
				−0.64			
TKP90					1.166		
					−0.7		
TKP105						1.908	
						−0.75	
TKP120							3.604
							−1.06
birthweight	−2.1E−05	−6.8E−06	−4.2E−06	−2.8E−05	−5.8E−05	−6.6E−05	−0.00011
	(−0.28)	(−0.11)	(−0.07)	(−0.34)	(−0.52)	(−0.57)	(−0.86)
i	0.514**	0.523***	0.518**	0.506**	0.458	0.474	0.355
	−3.08	−3.36	−3.24	−2.96	−1.82	−1.8	−0.75
icyst	0.654*	0.767***	0.807***	0.832***	0.822***	0.844***	0.71
	−2.03	−5.02	−6.82	−7.2	−4.68	−4.05	−1.27
ii	0.487*	0.538**	0.550**	0.516*	0.472	0.486	0.483
	−2.09	−2.66	−2.74	−2.33	−1.71	−1.65	−1.2
cholangitis	−0.181*	−0.194***	−0.203***	−0.202***	−0.185*	−0.162	−0.144
	(−2.57)	(−3.43)	(−3.77)	(−3.66)	(−2.34)	(−1.41)	(−0.85)
ass_anomalies	−0.186	−0.206	−0.203	−0.251	−0.227	−0.18	−0.153
	(−1.49)	(−1.60)	(−1.58)	(−1.60)	(−1.77)	(−1.37)	(−0.89)
HIC	0.158*	0.156*	0.161*	0.146*	0.123	0.136	0.111
	−2.43	−2.39	−2.4	−2.18	−1.38	−1.63	−0.85
dehycol	0.0887	0.103	0.0951	0.0899	0.0713	0.0889	0.0753
	−1.54	−1.92	−1.76	−1.6	−0.95	−1.39	−0.85
ursodeoxycholic acid	0.356***	0.354***	0.348***	0.349***	0.320*	0.319*	0.229
	−4.03	−4.05	−3.94	−3.86	−2.36	−2.03	−0.73
steroid	−0.242*	−0.212*	−0.205*	−0.186	−0.213*	−0.220*	−0.266*
	(−2.33)	(−2.19)	(−2.11)	(−1.75)	(−2.14)	(−2.16)	(−2.48)
_cons	0.127	0.053	0.0295	−0.137	−0.673	−1.44	−2.914
	−0.59	−0.26	−0.14	(−0.35)	(−0.61)	(−0.70)	(−1.02)
N	2405	2405	2405	2405	2405	2405	2405
t statistics in parentheses							
=″* p<0.05	** p<0.01	*** p<0.001″					

REFERENCES

Matsui, Akira, et al. "Direct Enzymatic Assay of Urinary Sulfated Bile Acids to Replace Serum Bilirubin Testing for Selective Screening of Neonatal Cholestasis." *Journal of Pediatric Surgery* 129 (1996): 306-8.

Alderman, Harold, et al. "Child Health and School Enrollment: A Longitudinal Analysis." *Journal of Human Resources* (2001): 185-205.

Alderman, Harold, John Hoddinott, and Bill Kinsey. "Long Term Consequences of Early Childhood Malnutrition." *Oxford Economic Papers* 58.3 (2006): 450-74.

Alderman, Harold, Hans Hoogeveen, and Mariacristina Rossi. "Preschool Nutrition and Subsequent Schooling Attainment: Longitudinal Evidence from Tanzania." *Economic Development and Cultural Change* 57.2 (2009): 239-60.

Angrist, Joshua D. "Treatment Effect Heterogeneity in Theory and Practice" *Economic Journal* 114.494 (2004): C52-C83.

Angrist, Joshua D, Guido W Imbens, and Donald B Rubin. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91.434 (1996): 444-55.

Angrist, Joshua D, and Victor Lavy. *Using Maimonides' Rule to Estimate the Effect of Class Size on Student Achievement*: National Bureau of Economic Research, 1999.

Angrist, Joshua David, and Jèorn-Steffen Pischke. *Mostly Harmless Econometrics : An Empiricist's Companion*. Princeton: Princeton University Press, 2009.

Wildhaber, Barbara E., et al. "Biliary Atresia: Swiss National Study, 1994–2004." *Journal of Pediatric Gastroenterology and Nutrition* 46 (2008): 299-307.

Bezerra, J. A., et al. "Use of Corticosteroids after Hepatoportoenterostomy for Bile Drainage in Infants with Biliary Atresia: The Start Randomized Clinical Trial." *Journal of the American Medical Association* 311.17 (2014): 1750-9.

Budtz-Jørgensen, Esben, et al. "Confounder Selection in Environmental Epidemiology: Assessment of Health Effects of Prenatal Mercury Exposure." *Annals of epidemiology* 17.1 (2007): 27-35.

Carvalho, Ed, et al. "Biliary Atresia: The Brazilian Experience." *Journal of Pediatrics* 86.6 (2010): 473-9.

Chardot, Christophe, et al. "Is the Kasai Operation Still Indicated in Children Older Than 3 Months Diagnosed with Biliary Atresia?" *Journal Pediatr* 138.2 (2001): 224-8.

Chardot, Christophe, et al. "Is the Kasai Operation Still Indicated in Children Older Than 3 Months Diagnosed with Biliary Atresia?" *The Journal of Pediatrics* 138.2 (2001): 224-28.

Chardot, Christophe, et al. "Improving Outcomes of Biliary Atresia: French National Series 1986-2009." *Journal of Hepatology* 58.6 (2013): 1209-17.

Chardot, Christophe, et al. "Prognosis of Biliary Atresia in the Era of Liver Transplantation: French National Study from 1986 to 1996." *Hepatology* 30.3 (1999): 606-11.

Chiu, C. Y., et al. "Biliary Atresia in Preterm Infants in Taiwan: A Nationwide Survey." *Journal of Pediatrics* 163.1 (2013): 100-3 e1.

Danks, D. M., et al. "Studies of the Aetiology of Neonatal Hepatitis and Biliary Atresia." *Archives of Disease in Childhood* 52.5 (1977): 360-7.

Davenport, M. "Biliary Atresia: Outcome and Management." *Indian Journal of Pediatrics* 73.9 (2006): 825-8.

Davenport, M., et al. "Biliary Atresia in England and Wales: Results of Centralization and New Benchmark." *Journal of Pediatric Surgery* 46.9 (2011): 1689-94.

De Vries, W., et al. "Biliary Atresia in the Netherlands: Outcome of Patients Diagnosed between 1987 and 2008." *Journal Pediatrics* 160.4 (2012): 638-44.

De Vries, W., et al. "Twenty-Year Transplant-Free Survival Rate among Patients with Biliary

Atresia." *Clinical Gastroenterol Hepatology* 9.12 (2011): 1086-91.

Dillon, P., et al. "Increased Expression of Intercellular Adhesion Molecules in Biliary Atresia." *American Journal of Pathology* 145.2 (1994): 263-7.

EURORDIS, European Organization for Rare Diseases. "Rare Diseases : Understanding This Public Health Priority." (2005).

Finn, Jeremy D, and Charles M Achilles. "Tennessee's Class Size Study: Findings, Implications, Misconceptions." *Educational Evaluation and Policy Analysis* 21.2 (1999): 97-109.

Fisher, Ronald A. "The Causes of Human Variability." *The Eugenics Review* 10.4 (1919): 213.

Frederick M. Karrer, John R. Lilly, Barbara A. Stewart, and Roberta J. Hall, and Colorado Denver. "Biliary Atresia Registry, 1976 to 1989." *Journal of Pediatric Surgery* 25.10 (1990): 1076-81.

Glewwe, Paul. "The Impact of Child Health and Nutrition on Education in Developing Countries: Theory, Econometric Issues, and Recent Empirical Evidence." *Food & Nutrition Bulletin* 26.Supplement 2 (2005): 235-50.

Glewwe, Paul, Hanan G Jacoby, and Elizabeth M King. "Early Childhood Nutrition and Academic Achievement: A Longitudinal Analysis." *Journal of Public Economics* 81.3 (2001): 345-68.

Glewwe, Paul, and Michael Kremer. "Schools, Teachers, and Education Outcomes in Developing Countries." *Handbook of the Economics of Education* 2 (2006): 945-1017.

Grantham-McGregor, S. "Early Child Development in Developing Countries." *Lancet* 369.9564 (2007): 824.

Hartley, J. L., M. Davenport, and D. A. Kelly. "Biliary Atresia." *Lancet* 374.9702 (2009): 1704-13.

Holland, Paul W. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81.396 (1986): 945-60.

Hsiao, C. H., et al. "Universal Screening for Biliary Atresia Using an Infant Stool Color Card in Taiwan." *Hepatology* 47.4 (2008): 1233-40.

Jimenez-Rivera, C., et al. "International Incidence and Outcomes of Biliary Atresia." *Journal of Pediatric Gastroenterology and Nutrition* 56.4 (2013): 344-54.

Jones, Laura L, et al. "A Comparison of the Socio-Economic Determinants of Growth Retardation in South African and Filipino Infants." *Public Health Nutrition* 11.12 (2008): 1220-28.

Krueger, Alan B, and Diane M Whitmore. "The Effect of Attending a Small Class in the Early Grades on College‐Test Taking and Middle School Test Results: Evidence from Project Star." *The Economic Journal* 111.468 (2001): 1-28.

Lien, T. H., et al. "Effects of the Infant Stool Color Card Screening Program on 5-Year Outcome of Biliary Atresia in Taiwan." *Hepatology* 53.1 (2011): 202-8.

Linnemayr, Sebastian, and Harold Alderman. "Almost Random: Evaluating a Large-Scale Randomized Nutrition Program in the Presence of Crossover." *Journal of Development Economics* 96.1 (2011): 106-14.

Luo, Yi, and Shan Zheng. "Current Concept About Postoperative Cholangitis in Biliary Atresia." *World Journal of Pediatrics* 4.1 (2008): 14-19.

Maldonado, George, and Sander Greenland. "Simulation Study of Confounder-Selection Strategies." *American journal of epidemiology* 138.11 (1993): 923-36.

Matsui A, Dodoriki M. "Screening for Biliary Atresia." *Lancet* 345 (1995): 1181.

Meyers, R. L., et al. "High-Dose Steroids, Ursodeoxycholic Acid, and Chronic Intravenous Antibiotics Improve Bile Flow after Kasai Procedure in Infants with Biliary Atresia." *Journal of Pediatric Surgery* 38.3 (2003): 406-11.

Muise, A. M., et al. "Biliary Atresia with Choledochal Cyst: Implications for Classification." *Clinical Gastroenterology and Hepatology* 4.11 (2006): 1411-4.

Muraji, T. "Early Detection of Biliary Atresia: Past, Present & Future." *Expert Review of Gastroenterology & Hepatology* 6.5 (2012): 583-9.

Murnane, Richard J., and John B. Willett. *Methods Matter : Improving Causal Inference in Educational and Social Science Research*. Oxford ; New York: Oxford University Press, 2011.

Nio, M., and T. Muraji. "Multicenter Randomized Trial of Postoperative Corticosteroid Therapy for Biliary Atresia." *Pediatric Surgery International* 29.11 (2013): 1091-5.

Nio, M., et al. "Long-Term Outcome in Type I Biliary Atresia." *Journal of Pediatric Surgery* 41.12 (2006): 1973-5.

Nio, M., et al. "Impact of Age at Kasai Operation on Short- and Long-Term Outcomes of Type Iii Biliary Atresia at a Single Institution." *Journal of Pediatric Surgery* 45.12 (2010): 2361-3.

Nio, Masaki, et al. "Five- and 10-Year Survival Rates after Surgery for Biliary Atresia: A Report from the Japanese Biliary Atresia Registry." *Journal of Pediatric Surgery* 38.7 (2003): 997-1000.

Oh, Michael, et al. "The Kasai Procedure in the Treatment of Biliary Atresia." *Journal of Pediatric Surgery* 30.7 (1995): 1077-81.

Pediatrics, American Academy of. "Child Population: Number of Children (in Millions) Ages 0–17 in the United States by Age, 1950–2013 and Projected 2014–2050." (2014).

Pollitt, Ernesto, et al. "Early Supplementary Feeding and Cognition: Effects over Two Decades." *Monographs of the Society for Research in Child Development* (1993): i-118.

Rubin, Donald B. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66.5 (1974): 688.

Schreiber, R. A., et al. "Biliary Atresia: The Canadian Experience." *Journal of Pediatrics* 151.6 (2007): 659-65.

Schreiber, R. A., et al. "Biliary Atresia in Canada: The Effect of Centre Caseload Experience on Outcome." *Journal of Pediatric Gastroenterology and Nutrition* 51.1 (2010): 61-5.

Schwarz, K. B., et al. "Extrahepatic Anomalies in Infants with Biliary Atresia: Results of a Large Prospective North American Multicenter Study." *Hepatology* 58.5 (2013): 1724-31.

Serinet, M. O., et al. "Impact of Age at Kasai Operation on Its Results in Late Childhood and Adolescence: A Rational Basis for Biliary Atresia Screening." *Pediatrics* 123.5 (2009): 1280-6.

Shneider, B. L., et al. "A Multicenter Study of the Outcome of Biliary Atresia in the United States, 1997 to 2000." *Journal of Pediatrics* 148.4 (2006): 467-74.

Sokol, R. J., et al. "Screening and Outcomes in Biliary Atresia: Summary of a National Institutes of Health Workshop." *Hepatology* 46.2 (2007): 566-81.

Stock, James H, Jonathan H Wright, and Motohiro Yogo. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business & Economic Statistics* 20.4 (2002).

Stringer, M. D., et al. "Kasai Portoenterostomy: 12-Year Experience with a Novel Adjuvant Therapy Regimen." *Journal of Pediatric Surgery* 42.8 (2007): 1324-8.

Superina, R., et al. "The Anatomic Pattern of Biliary Atresia Identified at Time of Kasai Hepatoportoenterostomy and Early Postoperative Clearance of Jaundice Are Significant Predictors of Transplant-Free Survival." *Ann Surg* 254.4 (2011): 577-85.

Superina, R., et al. "The Anatomic Pattern of Biliary Atresia Identified at Time of Kasai Hepatoportoenterostomy and Early Postoperative Clearance of Jaundice Are Significant Predictors of Transplant-Free Survival." *Annals of Surgery* 254.4 (2011): 577-85.

The, N. S., et al. "Risk Factors for Isolated Biliary Atresia, National Birth Defects Prevention Study, 1997-2002." *American Journal of Medical Genetics Part A* 143a.19 (2007): 2274-84.

Tiao, M. M., et al. "Epidemiological Features of Biliary Atresia in Taiwan, a National Study 1996-2003." *Journal of Gastroenterology and Hepatology* 23.1 (2008): 62-6.

Victora, Cesar G, et al. "Maternal and Child Undernutrition: Consequences for Adult Health and Human Capital." *Lancet* 371.9609 (2008): 340-57.

Wada, H., et al. "Insignificant Seasonal and Geographical Variation in Incidence of Biliary Atresia in Japan: A Regional Survey of over 20 Years." *Journal of Pediatric Surgery* 42.12 (2007): 2090-2.

Wildhaber, B. E. "Biliary Atresia: 50 Years after the First Kasai." *International Scholarly Research Notices Surgery* 2012 (2012)

Wildhaber, B. E., et al. "Biliary Atresia: Swiss National Study, 1994-2004." *Journal of Pediatric Gastroenterology and Nutrition* 46.3 (2008): 299-307.

Yoon, P. W., et al. "Epidemiology of Biliary Atresia: A Population-Based Study." *Pediatrics* 99.3 (1997): 376-82.

Zweig, Mark H, and Gregory Campbell. "Receiver-Operating Characteristic (Roc) Plots: A Fundamental Evaluation Tool in Clinical Medicine." *Clinical chemistry* 39.4 (1993): 561-77.