Useful Predictors of Bile Duct Stones in Patients Undergoing Laparoscopic Cholecystectomy

Alan N. Barkun, M.D., Jeffrey S. Barkun, M.D., Gerald M. Fried, M.D., Gabriella Ghitulescu, M.D., Oren Steinmetz, M.D., Christine Pham, M.D., Jonathan L. Meakins, M.D., D.Sc., Carl A. Goresky, M.D., Ph.D., and the McGill Gallstone Treatment Group

From the Divisions of Gastroenterology, General Surgery, and the Department of Radiology, McGill University, Montreal, Quebec, Canada

Objective
The authors determined the most useful predictors of common bile duct (CBD) stones as diagnosed by endoscopic retrograde cholangiopancreatography (ERCP) in patients who underwent laparoscopic cholecystectomy (LC).

Methods
Prospective and retrospective collection of historical, biochemical and ultrasonographic data was used. Receiver operating characteristics curve analysis was used to determine optimal biochemical cut-off values. Multivariate analysis using logistic regression with generation of the best model identifying independent predictors of CBD stones also was employed. Prospective validation of the model was performed on an independent group of patients.

Results
Endoscopic retrograde cholangiopancreatographies were performed before LC in 106 patients, and after LC in 33. Only four of ten clinical variables evaluated independently predicted the presence of CBD stones. The optimal model predicted a 94% probability of CBD stones in a patient older than 55 years of age who presented with an elevated bilirubin (over 30 µmol/L) and positive ultrasound findings (a dilated CBD, and a CBD stone seen on ultrasound). This model was validated prospectively in a subsequent series of 49 patients in which the probability of CBD stone was only 8% when all four predictors were absent.

Conclusions
The identified independent clinical predictors of a CBD stone helps select a population of symptomatic gallstone bearers who benefit most from cholangiographic assessment.

Laparoscopic cholecystectomy (LC) has become the new therapeutic gold standard in uncomplicated symptomatic cholelithiasis, and at least 80% of all patients are now treated this way. However, 3% to 33% of all patients with symptomatic gallstones may bear associated common bile duct (CBD) stones. The optimal approach to


Address reprint requests to Alan Barkun, M.D., Room D7.148, The Montreal General Hospital, Montreal, Quebec, Canada, H3G 1A4. Accepted for publication August 12, 1993.
patients with suspected choledocholithiasis undergoing LC remains uncertain and may vary according to local expertise. Endoscopic retrograde cholangiopancreatography (ERCP) before LC should not be performed systematically on all patients because of the associated morbidity and low yield. Further scrutiny of the usefulness of CBD stone predictors is, therefore, timely. In June 1990, when LC was first introduced in the McGill University teaching institutions, a common approach to patients scheduled to undergo LC was adopted i.e., all patients suspected of having a CBD stone were referred for preoperative ERCP, with attempts at endoscopic sphincterotomy when choledocholithiasis was confirmed. This practice permitted us to study the usefulness of different clinical predictors of CBD stones diagnosed by ERCP in a homogenous patient population. An optimal model of CBD stone prediction was determined by logistic regression and validated prospectively in a subsequent series of patients undergoing LC.

METHODS

Patient Population

The study population consisted of patients who were scheduled for or underwent LC, in whom ERCP was performed at one of four McGill University teaching hospitals between June 1990 and February 1992. Preoperative ERCP was performed if bile duct stones were suspected. Indications for ERCP included a history of jaundice or pancreatitis; elevations in serum bilirubin, alkaline phosphatase, gamma glutamyl transferase, aminotransferases, and amylase; and ultrasound findings of a dilated CBD or suspicion of a CBD stone. The criteria used for performing ERCPs after LC were similar to those adopted in the preoperative setting, but also included the intraoperative findings of a dilated cystic duct or CBD stone. In addition to these, postoperative indications for ERCP included the clinical suspicion of a bile leak or bile duct injury. Although not an indication for ERCP, ultrasonographic detection of multiple small gallbladder stones was noted and analyzed as a possible predictor of CBD stones.

Data Collection

All data were obtained from prospective surgical and endoscopic databases. When necessary, data also were collected prospectively and retrospectively from hospital charts, endoscopic reports, abdominal ultrasound, and ERCP films. Age, gender, and a past history of jaundice or pancreatitis were recorded. Biochemistry results were considered abnormal for a patient with any of the following: bilirubin > 17 µmol/L, alkaline phosphatase > 110 units/L, gamma glutamyl transferase (GGT) > 55 units/L, aminotransferases (aspartate transaminase [AST], alanine transaminase [ALT]) > 40 units/L, and amylase > 185 units/L. The highest serum levels within the 10 days before the ERCP examination were recorded for each patient and used for the analysis. Ultrasound studies were performed before ERCP. The common bile duct diameter (in millimeters) was measured in its mid-portion. In addition, a subjective assessment by a blinded examiner was used to determine whether or not there was CBD dilatation considering only the patient’s age and the overall appearance of the biliary tree. The suspicion or presence of bile duct stones seen on ultrasound was recorded as was the presence of small, numerous gallbladder stones. The ultrasound reports and films were subsequently reviewed and standardized by an investigator (CP) blinded to the results of the ERCP examinations. Endoscopic retrograde cholangiopancreatography findings, with respect to the presence of bile duct stones and other biliopancreatic pathologies, were reported by a certified radiologist. Endoscopic retrograde cholangiopancreatography films and reports were later reviewed independently by a blinded investigator (GG).

Statistical Analysis

The corresponding proportion and total number of patients in whom the information was available were expressed for each clinical predictor. The unit of analysis was a patient, not an ERCP examination. For patients who had more than one ERCP, only the variables noted before the patient’s first ERCP and the diagnosis at the definitive ERCP were used for the analysis. When distributed normally, continuous variables were described using mean and standard deviation; if not, they were expressed as median and range. Age was analyzed both as a continuous and a categorical variable using the observed median of the patient population as a cut-off. Test characteristics were determined for all clinical parameters, including sensitivity, specificity, and positive and negative predictive values. The ability of a biochemical test alone to differentiate between the absence and presence of a bile duct stone was assessed using receiver operating characteristics (ROC) curve analysis. An ROC curve displays the true-positive rate on the Y axis and the false-positive rate on the X axis for varying test thresholds, thus plotting the performance of a diagnostic test. For each biochemical test, five possible cut-offs were chosen after examination of its observed quartile values in the study population. The ideal cut-off for each biochemical test was chosen by determining the point closest to an ideal test with 100% specificity and sensitivity (the upper left most corner of the graph). Thereafter, the biochemical tests were analyzed in multivariate analysis using the
ideal cut-offs determined at ROC analysis. Common bile duct diameter at ultrasonography was analyzed categorically using the subjective assessment of enlargement and a cut-off value of >6 mm,\(^{14,15}\) with no difference in results. Chi square testing, or Fisher’s exact test (where appropriate),\(^6\) were used to determine those clinical variables that were associated with the presence of a CBD stone, when examined individually. Multiple stepwise logistic regression then was used to construct the model that would best predict the presence of a CBD stone at ERCP, considering all clinical predictors together. Goodness of fit of the model was assessed using the Pearson’s chi square test.\(^{15}\) Then a probability tree was developed to show the probability of CBD stones according to the presence or absence of the independent predictors identified. Prospective validation of the logistic regression model was performed on a subsequent series of patients scheduled for LC—who first underwent ERCP—by comparing their observed CBD stone prevalence with that estimated by the model using chi square testing.\(^{17}\) In addition, the parameter coefficients of the initial model were compared with those obtained after adding the data of the subsequent series of patients.

**RESULTS**

**Patient Characteristics**

The prospective McGill Laparoscopic Cholecystectomy Registry identified 1300 patients undergoing LC at four McGill University teaching hospitals during the 18-month study period.\(^{10}\) During this time, 127 ERCP examinations were performed before LC in 106 patients (mean = 1.2 ERCP per patient). After LC, 49 ERCPs were performed in 33 patients (mean = 1.5 ERCP per patient). All patients who underwent ERCP examinations were included in the analysis. The cohort mean age was 54.9 \(\pm\) 17.7 years (range = 17–91 years), and 63% of the study population were women (87 patients). Two patients required ERCP examinations before and after surgery (preoperative ERCP examinations were unsuccessful in one patient, and falsely negative in the other). During the study period, the intraoperative cholangiography rate was 4%.\(^{10}\) Further details on the outcome of patients who underwent intraoperative cholangiography are discussed elsewhere.\(^{10}\)

**Description of Clinical Predictors**

The mean ages were 57.5 \(\pm\) 16.8 years in the preoperative and 48.1 \(\pm\) 15.2 years in the postoperative patient populations. The frequency of the different predictors both preoperatively and postoperatively are displayed in

<table>
<thead>
<tr>
<th>Clinical Predictor</th>
<th>Preoperative</th>
<th>Postoperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (&gt;55 yrs)</td>
<td>66 (63%)</td>
<td>12 (36%)</td>
</tr>
<tr>
<td>Female Gender</td>
<td>66 (62%)</td>
<td>21 (64%)</td>
</tr>
<tr>
<td>History of jaundice</td>
<td>42 (41%)</td>
<td>8 (26%)</td>
</tr>
<tr>
<td>History of pancreatitis</td>
<td>18 (12%)</td>
<td>4 (14%)</td>
</tr>
<tr>
<td>Bilirubin &gt; 17 (\mu)mol/L</td>
<td>33 (33%)</td>
<td>15 (56%)</td>
</tr>
<tr>
<td>Alkaline phosphatase &gt; 110 units/L</td>
<td>69 (75%)</td>
<td>16 (62%)</td>
</tr>
<tr>
<td>GGT &gt; 55 units/L</td>
<td>47 (85%)</td>
<td>16 (84%)</td>
</tr>
<tr>
<td>AST &gt; 40 units/L</td>
<td>75 (77%)</td>
<td>21 (81%)</td>
</tr>
<tr>
<td>ALT &gt; 40 units/L</td>
<td>74 (77%)</td>
<td>20 (80%)</td>
</tr>
<tr>
<td>Amylase &gt; 185 units/L</td>
<td>24 (28%)</td>
<td>17 (68%)</td>
</tr>
<tr>
<td>Dilated CBD on U/S</td>
<td>39 (39%)</td>
<td>10 (38%)</td>
</tr>
<tr>
<td>CBD stone on U/S</td>
<td>22 (22%)</td>
<td>7 (25%)</td>
</tr>
<tr>
<td>Multiple, small gallbladder stones</td>
<td>77 (80%)</td>
<td>—</td>
</tr>
</tbody>
</table>

GGT = gamma glutamyl transferase; AST = Aspartate aminotransferase; ALT = Alanine aminotransferase; CBD = common bile duct; U/S = abdominal ultrasonography.

Dilated CBD on ultrasound was reported as such or over 6 mm in diameter.

Table 1. Postoperative indications for ERCP included a suspected retained stone in 25 patients (76%), and suspicion of a bile leak or CBD injury in 8 (24%). Mean elevations of biochemical predictors were 96 \(\pm\) 130 \(\mu\)mol/L for bilirubin, 243 \(\pm\) 143 units/L for alkaline phosphatase, 307 \(\pm\) 185 units/L for gamma glutamyl transferase, 258 \(\pm\) 261 units/L for AST, 308 \(\pm\) 264 units/L for ALT, and 1484 \(\pm\) 1841 units/L for amylase. The intraoperative finding of a dilated cystic duct could not be analyzed as a post-LC indication for ERCP because of missing or incomplete data. The time interval between the laboratory or imaging results and the initial ERCP examination ranged from 1 to 10 days for all patients.

**ERCP Findings**

Common bile duct stones were found in 50 of 106 (47%) patients preoperatively and 17 of 33 (52%) postoperatively (including 15 of the 25 [60%] patients in whom CBD stones were suspected post-LC). Endoscopic retrograde cholangiopancreatography did not alter patient management. Among the preoperative ERCPs, additional pancreaticobiliary information was obtained in five patients (15%). Pancreas divisum was diagnosed in three patients, and unsuspected chronic pancreatitis changes were found in three. In one of the patients with pancreas divisum, changes of chronic pancreatitis were found in the dorsal duct. Postoperative ERCP diagnosis of leaks and CBD injury are detailed elsewhere.\(^{18}\)
Determination of Optimal Serum Cut-Off Values for Biochemical Predictors

Using ROC curve analysis, optimal predictive serum cut-off levels were determined to be 30 $\mu$mol/L for bilirubin (Fig. 1), 500 units/L for amylase, 120 units/L for AST (similar results were found for ALT, which was available in fewer patients and, therefore, not studied further), and 300 units/L for alkaline phosphatase (Fig. 2) (GGT was not studied further because it was missing in many patients because the biochemical laboratories of participating institutions measure this serum value only if the alkaline phosphatase is elevated).

Accuracy of the Clinical Predictors

Table 2 details the performance of each possible predictor of CBD stone analyzed, including individual sensitivity, specificity, and positive and negative predictive values. When considered individually, the significant predictors of cholelithiasis included hyperbilirubinemia, a dilated CBD on ultrasonography, and the ultrasonographic suspicion of a CBD stone ($p < 0.01$). A history of pancreatitis and hyperamylasemia were associated significantly with the absence of a CBD stone.

Determination of the Multivariate Model

Using logistic regression, the best model for predicting bile duct stones at ERCP included the following independent predictors: age (continuous or dichotomized, using a cut-off of 55 years), an elevated bilirubin (over 30 $\mu$mol/L), the presence of a dilated CBD on ultrasonography ($> 6$ mm), and suspected or detected bile duct stone (at ultrasonography). The respective coefficients for each of these were 0.02 (0.88 when dichotomized), 0.83, 0.86, and 1.80. The choice of predictors did not change, whether the post-LC patients were excluded or included in the analysis. The two ultrasonographic predictors exhibited the largest coefficients and thus, were the strongest predictors of cholelithiasis. Depending on the presence or absence of these four parameters, the model yielded probabilities of finding a bile duct stone at ERCP that ranged from 18% (no predictor present), to 94% (all four predictors present). When a stone was seen on ultrasound, the overall probability of finding a stone at ERCP ranged from 58% to 94%, depending on the presence or absence of the other criteria. Figure 3 shows the set of probabilities expected according to the presence or absence of clinical predictors for patients in whom no CBD stone was noted on ultrasound.

Prospective Validation of the Model That Predicts Cholelithiasis at ERCP

The multivariate model was validated prospectively in a subsequent consecutive series of 49 patients suspected of having CBD stones and scheduled for LC. Common
bile duct stones were found at ERCP in only 1 of 12 (8%) patients in whom none of the four predictors were present, in contrast to 10 of 15 (66%) in whom two or more predictors were present. There were no significant differences between the observed and predicted probabilities in the subgroups of this patient population when broken down according to the different possible predictor combinations. In addition, the parameter coefficients of the model did not vary significantly when the data from the initial patients and subsequent 49 patients were combined.

**DISCUSSION**

The advent of LC has, rightly or wrongly, rekindled the debate of the optimal management of patients with suspected CBD stones who undergo cholecystectomy.\(^{19,20}\) One reason has been the desire to maintain the benefits of minimally invasive surgery. In addition, laparoscopic techniques to clear the CBD are evolving\(^ {21}\) but require a commitment of time, resources, and specialized equipment that is not widely available at present. The morbidity and success of many of these techniques have not been defined clearly or compared in large series to endoscopic sphincterotomy (before, during, or after LC) or open common bile duct exploration.\(^ {22-25}\) Predictors of CBD stones previously had been studied extensively in the era of open cholecystectomy. The reports, however, remained disparate in their conclusions, with different sets of predictors performing variably.\(^ {30}\) This inconsistency can be attributed to many factors, including differences in patient selection, differences in predictors studied and their timing with respect to the diagnosis of CBD stones, differences in gold standards used to detect the presence of choledocholithiasis, and the use of poorly adapted statistical techniques to what are often highly intercorrelated clinical predictors.
Some authors examined only preoperative predictors, and others examined a combination of preoperative and intraoperative variables. Furthermore, many of the studies were carried out before the availability of modern ultrasonographic equipment.

All studies were performed before the era of LC—perhaps an important consideration when taking into account the recently reported prevalence of CBD stones that appears to have dropped since the advent of LC.\textsuperscript{31,32} This decline may be real, possibly reflecting a different patient population that presents earlier in the course of symptomatic gallstone disease, or may be artificial and attributable to decreased detection.\textsuperscript{31}

Common bile duct stones are found in 3% to 9% of gallstone patients who have never been jaundiced.\textsuperscript{20,33–35} No patient with normal liver function tests was found with choledocholithiasis in two series totalling over 400 patients using intraoperative cholangiography (IOC) performed on patients at low risk for carrying CBD stones.\textsuperscript{36,37} Saltztein et al. demonstrated that only the combination of an elevated bilirubin and alkaline phosphatase within the 14 days before open cholecystectomy would significantly predict a CBD stone.\textsuperscript{38} Many studies have found ultrasonographic criteria to be necessary for the generation of a good predictive model\textsuperscript{39–41} because of their stronger predictive abilities when compared with other preoperative clinical criteria.\textsuperscript{41} Although suggestive of choledocholithiasis,\textsuperscript{42} a dilated CBD alone may be an unreliable indicator,\textsuperscript{43} perhaps because of the ability of the CBD diameter to change over a short period of time.\textsuperscript{44,45} Combining the CBD diameter with other abnormalities, such as an elevated bilirubin, increases its positive predictive value, perhaps by reflecting ongoing biliary obstruction.\textsuperscript{15} Conversely, CBD diameter less than 3 mm almost rules out a CBD stone.\textsuperscript{46} Lacaine et al. noted that a CBD diameter over 12 mm, with a raised alkaline phosphatase, minimally raised bilirubin, or history of jaundice gave a 90% to 100% predictive value of CBD stones.\textsuperscript{39} In their series, patients with none of these symptoms exhibited a less than 5% risk of choledocholithiasis. They also reported, as others have,\textsuperscript{15,41} a significantly older mean age in the patients with choledocholithiasis, but did not otherwise examine this finding. Limitations of this study included the exclusion of all patients with bilirubin levels greater than 3 mg/dL, the absence of multivariate modelling, and the unknown error for each subgroup point estimate. Wilson et al., using IOC, suggested a similar predictive index that included elevated levels of alkaline phosphatase or bilirubin in the past 6 months, an abnormal CBD, and a dilated cystic duct at surgery.\textsuperscript{43} Taylor et al., using multivariate analysis, correctly classified 90% of patients when considering only a dilated CBD at intraoperative cholangiography and the presence of numerous small gallbladder stones from a list of 36 possible preoperative and intraoperative variables.\textsuperscript{41} Hauer-Jensen et al. observed 393 patients for 6 to 8 years after surgery.\textsuperscript{40} Logistic regression recognized 5 of 11 possible clinical criteria as independent, significant predictors. The three preoperative variables were age, hyperbilirubinemia, and the presence of a CBD stone on preoperative imaging. A dilated cystic duct and a palpable CBD stone were the two intraoperative predictors also identified. However, no prospective validation of the model was carried out. Alkaline phosphatase did not add independent preoperative prediction beyond that of bilirubin and ultrasonography, yet the authors chose to keep this variable in their model because of other groups’ findings.\textsuperscript{39,41} Positive and negative predictive values of the model were 28% to 36% and 97% to 100%, respectively.

When interpreting the performance of the models developed, it must be remembered that intraoperative criteria carry the strongest predictive abilities.\textsuperscript{41,42,47,48} Unfortunately, their availability may have decreased in the era of LC (inability to palpate a CBD stone at laparoscopy). The model generated in the present study, which examines no intraoperative variables, includes all three preoperative predictors determined in the study by Hauer-Jensen\textsuperscript{40} and ultrasonographic CBD diameter, found by many groups to be useful.\textsuperscript{39–41} The superior predictive abilities of ultrasonographic criteria also are confirmed.\textsuperscript{41} Hyperamylasemia and a history of pancreatitis were not significant predictors of a CBD stone, as reported previously.\textsuperscript{15,40,49} Some have suggested further that hyperamylasemia may lower the predictive ability of other tests.\textsuperscript{38} Clinical series suggest that unsustained hyperamylasemia reflects recent stone migration across the sphincter of Oddi.\textsuperscript{50,51} Moreover, patients with pancreatitis are at high risk for further episodes during the next few months.\textsuperscript{52,53} Thus, hyperamylasemia may best be used as a predictor of recurrent pancreatitis rather than as a predictor of CBD stone discovery.\textsuperscript{52,53} In patients presenting with symptoms other than pancreatitis, the timing of laboratory findings in relationship to cholangiography and surgery remains poorly studied. Nonetheless, it is plausible that the shorter the time interval between biochemical abnormalities and cholangiography, the higher the positive prediction while the likelihood of subsequent stone migration across the sphincter of Oddi (a cause for false-positive prediction) decreases. The time intervals in our study between the onset of acute symptoms, laboratory data sampling, and ERCP (1–10 days) reflect real life conditions and make the observed predictors’ performance useful, as reported. The potential for further gallbladder stone migration is present, however, and consideration must be given to the optimal timing of LC after a negative ERCP. Certain groups have recommended that this interval not exceed
24 to 48 hours. These suggestions are based on a small number of anecdotal observations, and our experience in 106 patients does not support such recommendations.

Strengths of the present study lie in the use of multivariate modelling and the assessment of a homogenous patient population undergoing LC. This may differ from other endoscopic series, which might also include the results of ERCP examinations performed in patients long after open cholecystectomy. Such a population subset may not be biologically similar with respect to CBD distensibility and enzyme elevation in response to transient or persistent obstruction by retained or recurrent stones. The present study results are limited by partial reliance on retrospective data and the short follow-up time. This may have resulted in an underestimation of the true negative predictive value because ERCP was not, and should not be performed routinely on all patients preoperatively. An approximation of the true-negative predictive value in the absence of long-term follow-up can be derived only from studies using routine intraoperative cholangiography. Indeed, only then does the study population include patients exhibiting no predictors whatsoever. The strength of the present model, therefore, lies in its ability to predict the presence of a CBD stone, not its absence.

Cotton et al. and others have proposed different diagnostic or therapeutic approaches for patients stratified according to varying likelihoods of carrying a CBD stone before LC. In the present series, based on widely used preoperative criteria, the overall positive predictive value of a preoperative ERCP approach was 47%, a performance comparable to those of other endoscopic series using similar selection criteria (32–61%). The model we have generated predicts a CBD stone with a probability ranging from 19% to 38%, in which none or only one criterion is present, to 49% to 94%, in which three or four criteria are fulfilled. A selection based on these predictors should help tailor an appropriate cholangiographic approach, given an individual's risk of CBD stone. The present study design does not permit the identification of an optimal method of cholangiography (ERCP pre-LC or post-LC vs. IOC) because it fails to examine other important considerations, such as the morbidity of the procedure, the implications of a false-positive cholangiographic examination, or a determination of the impact of undetected CBD stones. More precise and unbiased quantifications of these associated factors, which may vary from center to center according to local expertise, are required urgently. Until then, based on this report and our previously published experience, we would recommend that centers with similar available expertise perform preoperative ERCP in patients exhibiting one or more of the aforementioned strong predictors (other than age > 55 years alone). Patients with “weaker” predictors might best be investigated with other methods, such as IOC. The lack of pertinent, comparable data underscores the importance of, and need for long-term prospective trials to better define the optimal roles of different approaches to patients suspected of choleodocholithiasis who undergo LC.

Acknowledgments

The authors thank Christine Wickham and Diane Thibeault for their assistance in data collection, and Mary Muccino for her assistance in the preparation of the manuscript.

References

17. Hosmer DW, Lemeshow S. Assessing the fit of the model. In