Original Investigation

Effect of Endoscopic Sphincterotomy for Suspected Sphincter of Oddi Dysfunction on Pain-Related Disability Following Cholecystectomy
The EPISOD Randomized Clinical Trial

Peter B. Cotton, MD, FRCP, FRCS; Valerie Durkalski, PhD; Joseph Romagnuolo, MD; Qi Pauls, MS; Evan Fogel, MD; Paul Tarnasky, MD; Giuseppe Aliperti, MD; Martin Freeman, MD; Richard Kozarek, MD; Priya Jamidar, MD; Mel Wilcox, MD; Jose Serrano, MD, PhD; Olga Brawman-Mintzer, MD; Grace Elta, MD; Patrick Mauldin, PhD; Andre Thornhill; Robert Hawes, MD; April Wood-Williams; Kyle Orrell; Douglas Drossman, MD; Patricia Robuck, PhD

IMPORTANCE Abdominal pain after cholecystectomy is common and may be attributed to sphincter of Oddi dysfunction. Management often involves endoscopic retrograde cholangiopancreatography (ERCP) with manometry and sphincterotomy.

OBJECTIVE To determine whether endoscopic sphincterotomy reduces pain and whether sphincter manometric pressure is predictive of pain relief.

DESIGN, SETTING, AND PATIENTS Multicenter, sham-controlled, randomized trial involving 214 patients with pain after cholecystectomy without significant abnormalities on imaging or laboratory studies, and no prior sphincter treatment or pancreatitis randomly assigned (August 6, 2008-March 23, 2012) to undergo sphincterotomy or sham therapy at 7 referral medical centers. One-year follow-up was blinded. The final follow-up visit was March 21, 2013.

INTERVENTIONS After ERCP, patients were randomized 2:1 to sphincterotomy (n = 141) or sham (n = 73) irrespective of manometry findings. Those randomized to sphincterotomy with elevated pancreatic sphincter pressures were randomized again (1:1) to biliary or to both biliary and pancreatic sphincterotomies. Seventy-two were entered into an observational study with conventional ERCP managemeny.

MAIN OUTCOMES AND MEASURES Success of treatment was defined as less than 6 days of disability due to pain in the prior 90 days both at months 9 and 12 after randomization, with no narcotic use and no further sphincter intervention.

RESULTS Twenty-seven patients (37%; 95% CI, 25.9%-48.1%) in the sham treatment group vs 32 (23%; 95% CI, 15.8%-29.6%) in the sphincterotomy group experienced successful treatment (adjusted risk difference, −15.6%; 95% CI, −28.0% to −3.3%; P = .01). Of the patients with pancreatic sphincter hypertension, 14 (30%; 95% CI, 16.7%-42.9%) who underwent dual sphincterotomy and 10 (20%; 95% CI, 8.7%-30.5%) who underwent biliary sphincterotomy alone experienced successful treatment. Thirty-seven treated patients (26%; 95% CI,19%-34%) and 25 patients (34%; 95% CI, 23%-45%) in the sham group underwent repeat ERCP interventions (P = .22). Manometry results were not associated with the outcome. No clinical subgroups appeared to benefit from sphincterotomy more than others. Pancreatitis occurred in 15 patients (11%) after primary sphincterotomies and in 11 patients (15%) in the sham group. Of the nonrandomized patients in the observational study group, 5 (24%; 95% CI, 6%-42%) who underwent biliary sphincterotomy, 12 (31%; 95% CI, 16%-45%) who underwent dual sphincterotomy, and 2 (17%; 95% CI, 0%-38%) who did not undergo sphincterotomy had successful treatment.

CONCLUSIONS AND RELEVANCE In patients with abdominal pain after cholecystectomy undergoing ERCP with manometry, sphincterotomy vs sham did not reduce disability due to pain. These findings do not support ERCP and sphincterotomy for these patients.

TRIAL REGISTRATION clinicaltrials.gov Identifier: 00688662

Supplemental content at jama.com

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Peter B. Cotton, MD, FRCP, FRCS, Digestive Disease Center, Medical University of South Carolina, ART 7094, MSC 290, 25 Courtenay Dr, Charleston, SC 29425 (cottonp@musc.edu).

Postcholecystectomy pain is a common clinical problem. More than 700 000 patients undergo cholecystectomy each year in the United States,[1] and at least 10% are reported to have pain afterwards.[2] A few are found by standard investigations to have a biliary cause (eg, duct stone) and some are diagnosed with other abdominal pathology or functional bowel disease.[3] Most have no significant abnormalities on imaging or laboratory testing, and the cause remains obscure. Many of these patients undergo endoscopic retrograde cholangiopancreatography (ERCP) in the hope of finding small stones or other pathology or in an effort to address suspected sphincter of Oddi dysfunction.[4] Of these patients, some undergo biliary or pancreatic sphincterotomy or both. The value of this endoscopic intervention is unproven and the risks are substantial. Procedure-related pancreatitis rates are 10% to 15%,[5] and perforations may occur. Many patients have prolonged and expensive hospital stays, and some die.[6]

Sphincter of Oddi dysfunction has been divided into 3 types.[3] Type I consists of patients with a dilated bile duct and abnormal liver tests, type II involves one of those criteria but not both, and type III have none of those criteria. A National Institutes of Health conference in 2002 raised concerns about the safety of ERCP in this context. It recommended that patients with suspected sphincter of Oddi dysfunction types II and III be referred to tertiary centers able to perform sphincter manometry.[7] However, sphincter manometry has never been shown to predict the outcome of sphincterotomy in patients with sphincter of Oddi dysfunction type III, and cohort studies have shown unimpressive results.[8-10] There has been only a single sham-controlled study that was reported in abstract form only, showing that 8 of 13 patients treated by biliary sphincterotomy improved compared with 3 of 10 control patients.[11] Placebo effects are likely strong; the pain response rates in the sham groups of the published trials in patients with sphincter of Oddi dysfunction type II were 33% and 42%.[12,13]

Thus, the current practice of performing ERCP in these patients, with or without sphincterotomy, and with or without manometry, is not supported by evidence. In addition to the significant risks of the procedure, many patients who do not benefit undergo more ERCPs and even surgical interventions. Several authors have expressed concerns about this practice.[8-10,14,15]

We performed a multicenter study to determine the effectiveness and safety of endoscopic sphincterotomy compared with sham treatment in adult patients with unexplained postcholecystectomy pain.

Methods

Study Design

The Evaluating Predictors and Interventions in Sphincter of Oddi Dysfunction (EPISOD) trial is a sham-controlled, randomized, double-blind clinical trial. Details of the design have been published.[16] The study protocol was approved by the institutional review board at each site, and all patients gave written informed consent.

Study Patients

Patients in 7 US centers were considered for the study if they were aged 18 to 65 years, had pain more than 3 months after cholecystectomy, no history of pancreatitis (or evidence for it on imaging and laboratory tests), and no prior sphincter intervention. Patients fulfilling these criteria were considered further if treatment with acid-suppressing agents and antispasmodics failed, had normal upper endoscopy and abdominal imaging (with a bile duct diameter of ≤9 mm), and if taking antidepressants were maintaining a stable dose of medications. They were excluded if within the last 6 months levels of direct bilirubin, alkaline phosphatase, amylase, or lipase were more than twice normal or if transaminases were more than 3 times normal; if narcotic analgesics were used every day in the past month; if they were known to have pancreas divisum, prior surgical biliary diversion, abnormal endoscopic ultrasound (if done), significant psychiatric or other major medical problems; or, were pregnant.

The proposed randomized study was explained to patients fulfilling the initial criteria. Those consenting completed a battery of questionnaires, including demographics, pain burden, details of the cholecystectomy and its effect on pain, psychological disorders as measured by the Mini International Neuropsychiatric Interview (MINI)[17]; Beck Depression Inventory (BDI-II)[18]; Hospital Anxiety and Depression Scale (HADS)[19]; Trauma Questionnaire-Short Form[20]; Coping Questionnaire-Catastrophizing Subscale[21]; Rome III Functional Gastrointestinal Disorders Diagnostic Module,[22] including the Biliary Disorders module, which was modified to allow daily pain (≤2 on a 0-10 scale) as well as intermittent episodes; 36-Item Short Form Health Survey (36-SH) for quality of life[23]; and economic resource use. The patient's pain disability was assessed using the Recurrent Abdominal Pain Intensity and Disability (RAPID) instrument, which was developed and validated specifically for this study[24] The RAPID score is the patient's recall of how much productivity in 3 domains (paid work or school, household activities, and nonwork activities) was lost in the prior 90 days due to abdominal pain episodes. Thus, the score ranges from 0 to 270. The RAPID score has 4 grades: grade 1 indicates 6 days or less; grade 2, 6 to 10 days; grade 3, 11 to 20 days; and grade 4, 20 days or more.[24]

Patients were excluded from the study if they reported major psychiatric disorders (psychotic or bipolar disorder), severe depression (BDI score, ≥22), suicidal risk, insufficient pain disability (RAPID score, <11 days of disability in the past 3 months), or pain characteristics not consistent with the modified Rome III definition. Two of the entry criteria were revised early in the study. The pain criterion was changed to allow daily discomfort as well as pain episodes (modifying the Rome III criteria). In addition, the acceptable limit for transaminases of twice normal was raised to 3 times normal. These changes occurred when only 10 patients had been randomized, with approval of the data safety and monitoring board.

Patients who met all of the above criteria but declined randomization were offered enrollment into an observational study, EPISOD 2. They were studied and followed up in the same manner as patients randomized to the EPISOD trial, ex-

cept that biliary, pancreatic sphincterotomies, or both were performed based on manometry results, as is commonly practiced, and the patients were not blinded.

Study Intervention

Eligible patients underwent ERCP by very experienced endoscopists, under conscious sedation, modified or full anesthesia, based on local practice. Sphincter of Oddi manometry was performed by the standard water-perfusion method, using a basal pressure of more than 40 mm Hg in both leads to define abnormality in the biliary and pancreatic sphincters. Using a centralized web-based system, patients with successful pancreatic manometry and not having duct abnormalities, such as pancreas divisum, were then randomized with a 2:1 allocation to sphincterotomy or sham. The randomization algorithm was a permuted block scheme with random block size stratified by clinical center and the presence or absence of pancreatic sphincter hypertension. Because experts in tertiary centers will sometimes treat patients with biliary-type pain and pancreatic sphincter hypertension with pancreatic sphincterotomy, a secondary aim was to compare the results of dual (biliary and pancreatic) sphincterotomy with those of biliary sphincterotomy alone. Thus, patients allocated to sphincterotomy who had pancreatic sphincter hypertension were randomized a second time using a 1:1 ratio to either biliary or dual sphincterotomy.

All patients (including those in the sham group with no sphincterotomy) received small-caliber (3-5 Fr diameter) pancreatic stents to reduce the risk of postprocedure pancreatitis; these pass spontaneously within 1 to 4 weeks. Antiinflammatory medications now known to reduce postprocedure pancreatitis were not given. Patients were blinded to treatment allocation, observed in the hospital overnight, and returned to their blinded referring physicians for clinical follow-up. Blinded research coordinators at each site called the patients 1 week after undergoing the procedure and monthly for 12 months. In addition, calls were made at 9 and 12 months by blinded research staff at the central coordinating center to collect primary outcome information. Patients with persistent burdensome pain either consulted their referring physicians, who decided on treatment blinded to the original treatment allocation, or returned to the study site and were seen by an evaluating physician, also blinded to the treatment allocation. By definition, treatment was considered failed if a patient had undergone a second ERCP intervention or if the patient was recommended by the evaluating physician to undergo a second intervention, even if the treating physician disagreed.

Study Outcomes

The primary outcome was a dichotomous (success/failure) variable. Treatment success was defined as patients having a RAPID score of fewer than 6 days of lost productivity at months 9 and 12 after the endoscopic procedure, did not undergo a second intervention, and did not require narcotics during months 10, 11, and 12 unless they needed it for other than abdominal pain for no more than 14 days. If the 12-month RAPID score was missing or collected outside the acceptable window (within 6 months of the expected 12-month visit), the treatment was con-

sidered to have failed the patient for the primary outcome. If the 9-month RAPID score was missing or was outside of the window (within 3 months of the expected 9-month visit), then the 6-month value was used instead when available.

Secondary outcomes included association of manometry results and prespecified potential prognostic clinical factors (age, time since cholecystectomy, whether stones were present, whether there was a period of pain relief after surgery, presence of daily discomfort, elevated liver enzymes, other concomitant functional digestive disorders, narcotic use, and psychiatric status) with the primary outcome, the outcome of patients receiving biliary sphincterotomy compared with dual sphincterotomy, quality of life, and resource utilization. Adverse events were collected throughout the study period and were monitored by an independent medical safety monitor, and a data and safety monitoring board.

Statistical Analysis

The trial was designed to test for an overall absolute difference of at least 30% in the primary outcome in patients treated with sphincterotomy (dual or biliary) compared with those treated with sham. The sham success rate was expected to be 30%.[8,9,12,13] The choice of an absolute difference between treatment groups of 30% was based on published data[7-9] and pilot studies conducted during the planning stages of this study.

Based on this information and the 2:1 allocation, an assumed 10% nonadherence rate, and 1 interim analysis for efficacy using O'Brien and Fleming boundaries[25] and futility using conditional power, the study required 214 patients to be randomized to ensure greater than 90% likelihood of identifying this difference.
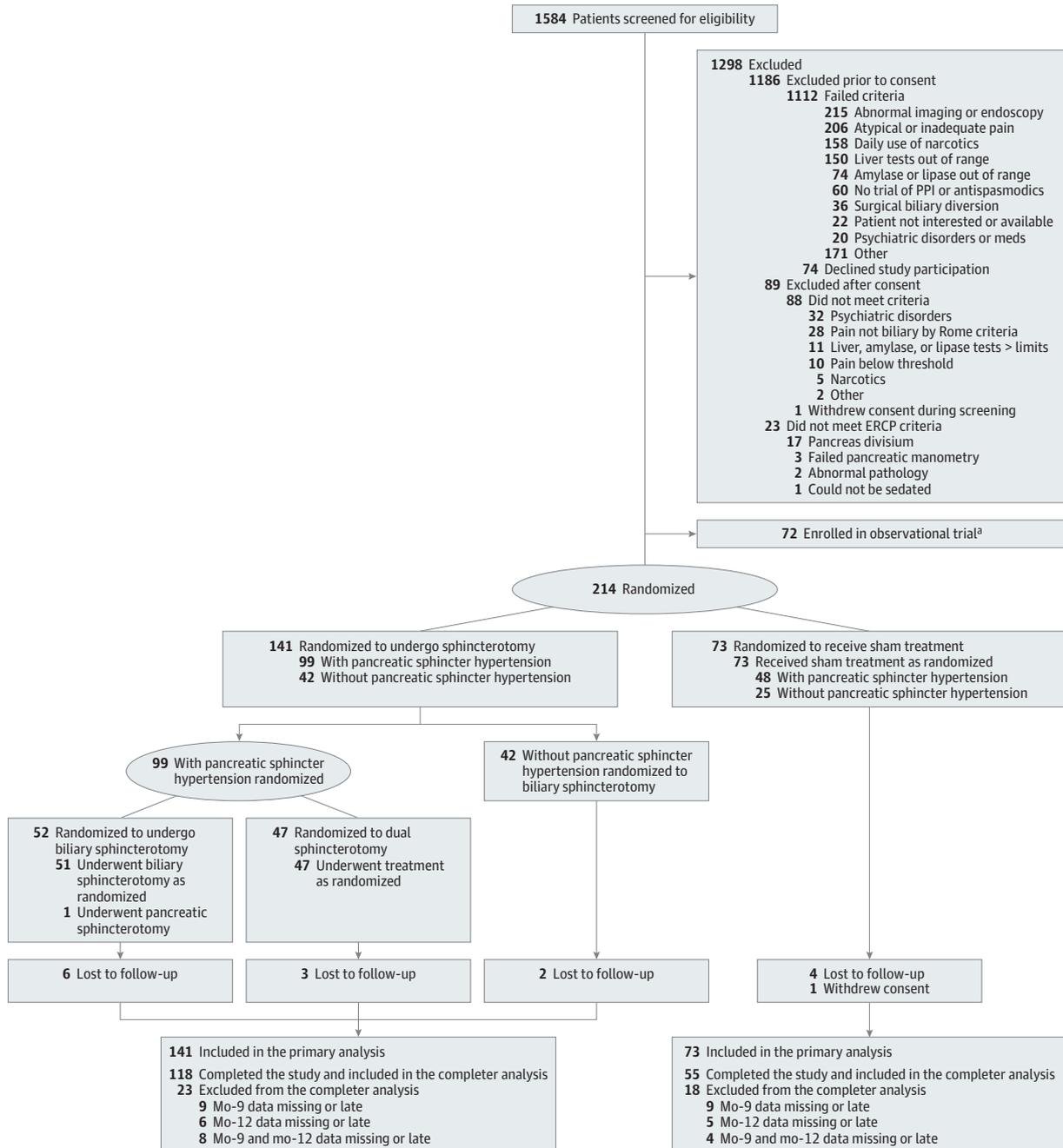
The primary analysis was conducted using a logistic regression model with the treatment group as the factor of interest and clinical center and pancreatic sphincter hypertension status as covariates. An unadjusted analysis of the primary outcome was conducted as a sensitivity analysis. A Wald test was used to compare the treatment group proportions using a 2-tailed significance level of .05. Adjusted and unadjusted risk differences with 2-sided 95% confidence intervals are reported. All analyses were conducted with the intention-to-treat population defined as all randomized patients. A sensitivity analysis was conducted with the completer population defined as all randomized patients who did not have a missing or late RAPID score at months 9 and 12.

Results

Patients and Baseline Characteristics

A total of 1584 fitting the 4 basic criteria (postcholecyctectomy pain, aged 18-65 years, no pancreatitis and no prior sphincter intervention) patients underwent preliminary screening prior to consent. Of these, 1112 (70%) were excluded. Many patients met more than 1 of the exclusion criteria. The following are the most common: 215 had abnormal findings on abdominal imaging or upper endoscopy; 206, atypical or little pain; 158, daily use of narcotics; and 224, liver or pancreas laboratory screening results higher than allowed limits. In addi-

Figure 1. Flow Diagram Documenting Recruitment and Treatment Allocations

```
                    ┌──────────────────────────────────────────┐
                    │ 1584  Patients screened for eligibility    │
                    └──────────────────────────────────────────┘
                                      │
                                      │    ┌─────────────────────────────────────────────────┐
                                      │    │ 1298 Excluded                                     │
                                      │    │   1186 Excluded prior to consent                  │
                                      │    │     1112 Failed criteria                          │
                                      ├───▶│        215 Abnormal imaging or endoscopy          │
                                      │    │        206 Atypical or inadequate pain            │
                                      │    │        158 Daily use of narcotics                 │
                                      │    │        150 Liver tests out of range               │
                                      │    │         74 Amylase or lipase out of range         │
                                      │    │         60 No trial of PPI or antispasmodics      │
                                      │    │         36 Surgical biliary diversion             │
                                      │    │         22 Patient not interested or available    │
                                      │    │         20 Psychiatric disorders or meds          │
                                      │    │        171 Other                                  │
                                      │    │       74 Declined study participation             │
                                      │    │     89 Excluded after consent                     │
                                      │    │       88 Did not meet criteria                    │
                                      │    │         32 Psychiatric disorders                  │
                                      │    │         28 Pain not biliary by Rome criteria      │
                                      │    │         11 Liver, amylase, or lipase tests > limits│
                                      │    │         10 Pain below threshold                   │
                                      │    │          5 Narcotics                              │
                                      │    │          2 Other                                  │
                                      │    │        1 Withdrew consent during screening        │
                                      │    │     23 Did not meet ERCP criteria                 │
                                      │    │       17 Pancreas divisium                        │
                                      │    │        3 Failed pancreatic manometry              │
                                      │    │        2 Abnormal pathology                       │
                                      │    │        1 Could not be sedated                     │
                                      │    └─────────────────────────────────────────────────┘
                                      │
                                      │    ┌─────────────────────────────────────────────────┐
                                      ├───▶│ 72 Enrolled in observational trial[a]             │
                                      │    └─────────────────────────────────────────────────┘
                                      │
                            ┌──────────────────┐
                            │  214  Randomized  │
                            └──────────────────┘
```


a Patients who were eligible for the study but who declined randomization were treated by conventional manometry-directed sphincterotomy. ERCP indicates
  endoscopic retrograde cholangiopancreatography; PPI, proton pump inhibitor.

tion, 74 patients (5%) declined consent. Three hundred ninety-eight patients consented and completed study questionnaires. Of those, 88 did not meet the specific criteria. The most common reasons were that 32 patients had established psychiatric disorders and 28 did not have biliary pain by modified Rome III criteria. One patient withdrew consent. Of the remaining 309 potential candidates, 72 declined randomization but con-

sented to be enrolled into the observational study. Two hundred thirty-seven patients underwent ERCP. Twenty-three of those patients were excluded: 17 had pancreas divisum, 2 did not have clinically relevant pathology, 3 had unsuccessful manometry, and 1 could not be adequately sedated (**Figure 1**).

Of the 214 patients who were randomized, 73 were assigned to the sham group and 141 to sphincterotomy group.

Table 1. Baseline Characteristics of Study Population by Assigned Treatment Group[a]

Demographics	No. (%) of Patients	
	Sphincterotomy (n = 141)	Sham (n = 73)
Age, mean (SD) [range], y	38 (11) [19-64]	39 (11) [20-59]
Women	128 (91)	69 (95)
Race/ethnicity		
Black	5 (4)	3 (4)
White	132 (94)	70 (96)
Other race/unknown	4 (2)	0 (0)
Employment		
Full-time	81 (57)	43 (59)
Part-time	20 (14)	12 (16)
Homemaker	17 (12)	4 (5)
Elevated liver enzymes[b]	11 (8)	8 (11)
Gallstones at cholecystectomy	67 (48)	34 (47)
Pain relief after cholecystectomy	100 (71)	47 (64)
Daily abdominal discomfort in past 30 d	74 (52)	36 (49)
Abnormal sphincter manometry	110 (78)	52 (71)
Pancreatic only	42 (30)	22 (30)
Biliary only	16 (11)	9 (12)
Pancreatic and biliary	52 (37)	21 (29)
Irritable bowel syndrome	45 (32)	28 (38)
DSM-IV depressive disorders (current)[c]	10 (7)	7 (10)
DSM-IV anxiety disorders (current)[c]	9 (6)	10 (14)
Physical/sexual abuse	38 (27)	13 (18)
Narcotics use (in last mo)	40 (28)	16 (22)
Antidepressant/anxiolytic use	55 (39)	29 (40)
Time from cholecystectomy to randomization, mean (SD) [range], y	4 (4.57) [0.27-21.08]	5 (6.67) [0.25-41.67]
Pain; Baseline RAPID score, mean (SD) [range][c]	82 (52.76) [11-251]	89 (67.28) [11-270]
Days of pain episodes in past 90 d, mean (SD) [range]	69 (27.08) [0-90]	69 (26.10) [3-90]
Average pain intensity (0-10) in past 90 d, mean (SD) [range]	7 (1.83) [3-10]	7 (1.99) [3-10]
SF-36 Pain Scale, mean (SD) [range][d]	35 (15.58) [0-84]	35 (18.26) [0-100]
HADS, mean (SD) [range][c]		
Anxiety (0-21)	4 (3.56) [0-16]	5 (3.37) [0-13]
Depression (0-21)	3 (3.12) [0-16]	4 (3.05) [0-11]
BDI-II (0-63), mean (SD) [range][c]	7.5 (5.32) [0-23]	8 (5.54) [0-26]
CSQ-CAT score (0-36) (coping/catastrophizing), mean (SD) [range][c]	7 (6.30) [0-32]	9 (7.26) [0-28]
SF-36, mean (SD) [range][d]		
Physical	38 (8.19) [19.43-54.69]	40 (7.22) [20.31-56.58]
Mental	49 (9.20) [17.86-66.89]	48 (10.36) [16.37-63.32]

Abbreviations: BDI, Beck Depression Inventory; CSQ-CAT, Coping Questionnaire Catastrophizing Subscale; *DSM-IV, Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition); HADS, Hospital Anxiety and Depression Scale; RAPID, Recurrent Abdominal Pain Intensity and Disability; SF-36, 36-Item Short Form Health Survey.

[a] There was no significant difference between the 2 treatment groups.

[b] Elevated transaminases of at least 3 × normal or alkaline phosphatase at least 2 × normal.

[c] Higher scores mean worse health.

[d] Higher score means better health.

Patients were randomized between August 6, 2008, and March 23, 2012. The last follow-up occurred on March 21, 2013. The 2 treatment groups were well balanced with respect to baseline demographic and clinical characteristics (**Table 1**), as described elsewhere.[26] Pancreatic sphincter manometry was abnormal (with or without biliary abnormality) in 64% of the patients; biliary manometry alone was abnormal in 12%. As reported elsewhere, abnormal sphincteric pressures were not influenced by the type of sedation or anesthesia and did not correlate with the patient's clinical characteristics.[27]

Primary Outcome

Most patients in both study groups showed considerable reductions in their RAPID score by 3 months (**Figure 2**). The interim analysis was conducted after one-third, 72 patients, completed the 12-month follow-up. At that time, the *P* value of a χ^2 test was .03, failing to reach the stopping boundary for overwhelming efficacy. The conditional power under the alternative hypothesis was 64% and 2% under the null hypothesis. Based on this information, the data and safety monitoring board recommended continuation of the trial. The rate of suc-

Figure 2. RAPID Score Distribution by Assigned Treatment Group and Visit

The boxes indicate interquartile ranges; circle within box, mean; horizontal line within box, median; error bars, 1.5 times the interquartile range; and circles, outliers. RAPID indicates Recurrent Abdominal Pain Intensity and Disability.
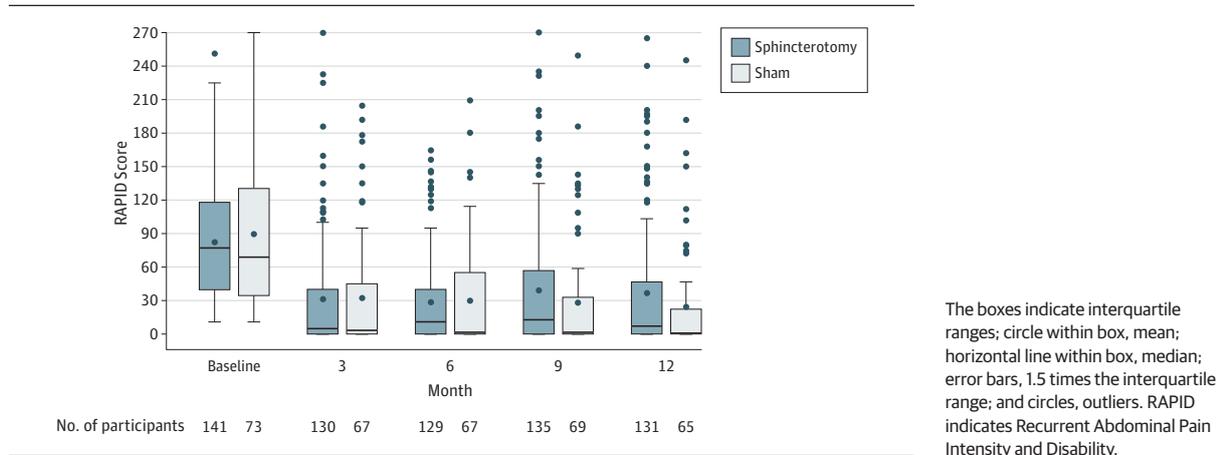
Table 2. Primary Outcome by Treatment Group and Success Rate for the Secondary Comparison of Biliary vs Dual Sphincterotomy for the Pancreatic Sphincter Hypertension Subgroup

| | Treatment | No. of Patients | No. (%) [95% CI] of Treatment Success | Risk Difference (95% CI) | |
				Adjusted[a]	Unadjusted
Primary outcome	Sham	73	27 (37) [21.6 to 33.6]	−15.7 (−28.0 to −3.3)	−14.3 (−27.3 to −1.2)
	Sphincterotomy (any)	141	32 (23) [15.8 to 29.6]		
Secondary outcome	Pancreatic sphincter hypertension with biliary sphincterotomy	51	10 (20) [8.7 to 30.5]		−10.2 (−27.2 to 6.8)
	Pancreatic sphincter hypertension with pancreatic and biliary sphincterotomy	47	14 (30) [16.7 to 42.9]		

[a] Sham vs sphincterotomy P value adjusted for site and pancreatic sphincter hypertension status is .01; unadjusted P value, .03.

cessful outcome as defined at 12 months was 37% for the 73 patients assigned to sham, and 23% for the 141 assigned to sphincterotomy (adjusted risk difference, −15.6%; 95% CI, −28.0% to −3.3%; P = .01; Table 2). A total of 23 patients had missing or late RAPID scores and no record of narcotic use or reintervention): 15 (11%) in the sphincterotomy group and 8 (11%) in the sham group. The primary outcome was imputed for these cases according to the study missing data plan described above.

The most common reason for failure in both treatment groups (72% sphincterotomy, 56% sham) was persistent elevation in the RAPID pain-disability score, with or without reintervention or narcotic use (eTable in the Supplement). A secondary sensitivity analysis of the primary outcome of the 173 patients that had complete outcome data revealed similar results. The success rate for the completer population in the sham group was 42% and in the sphincterotomy group, it was 25% (adjusted risk difference, −18.6%; 95% CI, −33.0% to 4.3%).

Safety Outcomes

The ERCP procedures at randomization caused pancreatitis in 26 patients, 11% in the sphincterotomy group and 15% in the sham group (unadjusted relative risk, 0.71; 95% CI, 0.34-1.46). Two of the events were defined as severe (1 in the biliary sphincterotomy group, 1 in sham group), 10 moderate, and

14 mild, using standard consensus criteria.[28] One patient in the dual group experienced retroduodenal perforation after sphincterotomy, which required surgical treatment. Another, in the sham group, had evidence for microperforation in association with pancreatitis. There were no procedure-related episodes of bleeding or infection and no deaths.

Secondary Outcomes

Ten patients 20% (95% CI, 8.7%-30.5%) with pancreatic sphincter hypertension assigned to the biliary sphincterotomy group and 14 patients 30% (95% CI, 16.7%-42.9%) assigned to dual biliary and pancreatic sphincterotomy experienced successful treatment (Table 2).

None of the prespecified clinical factors predicted success; these included age, reason for cholecystectomy and response to it, pain characteristics, and psychosocial comorbidities. In particular, there was no difference in outcomes between patients with and without daily abdominal discomfort or between those with and without any minor liver test abnormalities.

To examine whether our definition of treatment success was robust, we considered other post hoc criteria. Although the success rates were higher when we substituted a 50% reduction in the RAPID score (still counting reinterventions and narcotic use as failures), there was no difference between groups: sham success, 47% (95% CI, 35.1%-58.0%); sphincter-

Table 3. Secondary Outcomes Month-12 Change From Baseline[a]

	Sphincterotomy			Sham		
	No.	Mean (95% CI)	Median	No.	Mean (95% CI)	Median
RAPID score	131	−45 (−56.0 to −33.9)	−40	65	−62 (−83.3 to −40.1)	−38
HADS						
Anxiety	131	−1 (−1.7 to −0.2)	−1	68	−1 (−2.2 to −0.4)	−1
Depression	131	−1 (−1.7 to −0.3)	−1	68	−1 (−2.0 to −0.3)	−1
SF-36						
Physical	130	7 (5.2 to 8.6)	7	68	8 (5.8 to 10.3)	8
Mental	130	3 (1.5 to 5.1)	3	68	4 (1.8 to 7.1)	4

Abbreviations: HADS, Hospital Anxiety and Depression scale; RAPID, Recurrent Abdominal Pain Intensity and Disability; SF-36, 36-Item Short Form Health Survey.

[a] There were no statistical differences between sphincterotomy and sham for each outcome; data include patients with reinterventions and narcotic use but exclude patients with missing outcome data at month 12.

Table 4. Manometry and Outcomes in Patients Assigned to Sphincterotomy and Sham Group[a]

Manometry Sphincter Pressures		Sphincterotomy				Sham	
		Biliary		Biliary and Pancreatic			
Biliary	Pancreatic	No. of Patients	No. (%) of Treatment Success	No. of Patients	No. (%) of Treatment Success	No. of Patients	No. (%) of Treatment Success
Abnormal	Abnormal	25	5 (20)	27	9 (33)	21	8 (38)
Normal or 0	Abnormal	25	5 (20)	17	4 (23)	22	8 (36)
Abnormal	Normal	14	3 (21)	2	0	9	1 (11)
Normal or 0	Normal	30	5 (17)	1	1 (100)	21	10 (48)

[a] Pancreatic sphincter hypertension is based on reported sphincter pressures. Three patients with normal pancreatic pressures received dual sphincterotomies, and 1 patient with abnormal pancreatic pressure received biliary sphincterotomy. Abnormal is defined as pressures higher than 40 mm Hg in both leads. Zero means biliary manometry was not performed.

otomy, 38% (95% CI, 29.6%-45.6%). Similar results were found when the narcotic criteria were removed: sham success, 38% (95% CI, 27.2%-49.5%); sphincterotomy, 27% (95% CI, 19.6%-34.3%). The reintervention rate, when examined alone, was not significantly higher in the sham group (34%; 95% CI, 23%-45% vs 26%; 95% CI, 19%-34%; P = .22). In a sensitivity analysis to assess the extreme effect of sphincterotomy, we considered treatment to be successful among all the patients with missing or late follow-up data vs considering treatment failure among patients in the sham group whose data were missing or late and found that the treatment success rate for the sphincterotomy group was 31.2% (95% CI, 23.6%-38.9%) vs 31.5% (95% CI, 20.8%-42.2%) for the sham group. The adjusted risk difference was −4.8%, (95% CI, −17.0% to 7.4%; P = .44, adjusted for pancreatic sphincter hypertension and study site).

The 12-month SF-36 composite scores for physical and mental health and the HADS anxiety and depression scores improved from baseline; however, the change scores did not differ between the 2 treatment groups (Table 3).

The association between the results of manometry and primary outcome in those randomized to sphincterotomy was examined. The presence of pancreatic sphincter hypertension (regardless of biliary hypertension) was not associated with the primary outcome (P = .70). In addition, patients with pancreatic or biliary hypertension did not have a higher success rate than those without sphincter hypertension (P = .55; Table 4).

We similarly found no association with the primary outcome when we explored higher manometric pressure cut points.

In the observational study of patients who were not randomized, the use of sphincterotomy was guided by the results of manometry, as is usually done in clinical practice. The baseline demographic and clinical characteristics of the 72 patients were similar to the randomized population.[26] The success rates for these unblinded patients, using the same criteria as in EPISOD, were also similar: 24% (n = 5; 95% CI, 6%-42%) for biliary sphincterotomy, 31% (n = 12; 95% CI, 16%-45%) for dual sphincterotomy, and 17% (n = 2; 95% CI, 0%-38%) for those with normal manometry and no sphincterotomy.

Discussion

We report the results of a randomized, sham-controlled trial of sphincterotomy involving patients with pain after cholecystectomy and without significant liver test abnormalities or a dilated bile duct. Although many patients had a considerable reduction in their disability due to pain, we found that sphincterotomy was not more effective than a sham endoscopic procedure and that manometric pressure findings were not associated with the sphincterotomy outcome. Furthermore, we did not find any clinical characteristics that were associated with success, including factors proposed previously, such as age, pancreatic manometric pressures,[29] and pain

patterns.[30] We confirmed that ERCP carries substantial risks, even in expert hands.

The finding that endoscopic sphincterotomy is not an effective treatment has major implications for clinical practice because it applies to many thousands of patients. Several series have reported residual or recurrent pain in more than 20% of patients after elective cholecystectomy.[2,31-34] The number of cholecystectomies has increased recently, especially for suspected gallbladder dyskinesia.[35] Furthermore, the US Householders Survey[36] found that 1.5% of the adult population reported symptoms consistent with sphincter of Oddi dysfunction.

Our conclusions apply to patients with sphincter of Oddi dysfunction type III characteristics. It is noteworthy that we observed no benefit for sphincterotomy in the small number of our patients who did have some liver test abnormalities (ie, type II). Application of manometry-directed biliary sphincterotomy in type II patients is based on limited data,[11,12] and further study of type II patients may be justified.

Several potential study limitations need to be considered. First, the primary outcome measure might be considered to be too restrictive. The study was designed intentionally to have a stringent primary outcome in view of the known risks of ERCP and sphincterotomy. However, the conclusion remained robust in post hoc analyses using less stringent diagnostic criteria. Other sphincterotomy studies showing favorable outcomes used subjective and not objective measures of success[7,8] or defined success only as the lack of further intervention.[37] A second potential criticism, that patients who consent to a randomized sham-controlled trial might not be representative of most patients with the condition, was addressed by examination of the patients who declined to be randomized and were enrolled in the observational study. These patient's basal characteristics and outcomes were similar to the patients who were randomized. Another potential limitation was that our protocol deviated slightly from the strict Rome III definition of biliary pain[3] by allowing enrollment of patients with some daily discomfort in addition to episodes of pain. However, the results were no different in the patients with daily discomfort, nor was it different in those who had some minor laboratory abnormalities. Lastly, patients in the sham group underwent ERCP cannulation, manometry, and temporary stenting. It could be argued that these procedures were themselves somehow therapeutic, but it is highly unlikely that any such effect would last for a year. We did not include a second no-touch sham group because that would have increased complexity and cost and may have adversely affected recruitment.

We speculate that the early reduction in pain-related disability in all groups may be related to a placebo response in a cohort of optimistic and distressed patients who received support by continuing contact with our research staff. Other surgical and endoscopic randomized studies demonstrated similar success rates of about 30% in sham groups,[38,39] with benefits persisting for years.[40] A blinded cohort of the EPISOD and EPISOD 2 patients is being followed up for an additional 2 years to further assess this phenomenon. We have also considered the potential for regression to the mean to explain the large decrease in RAPID scores from baseline. Although that is possible, the reduction occurred equally in all study groups and does not affect the overall conclusions of the trial.

Conclusions

Among patients with abdominal pain after cholecystectomy (and with no significant laboratory or imaging abnormalities) undergoing ERCP with manometry, sphincterotomy compared with a sham procedure did not reduce disability due to pain. These findings do not support the use of ERCP and sphincterotomy for these patients.

REFERENCES

1. Duncan CB, Riall TS. Evidence-based current surgical practice: calculous gallbladder disease. *J Gastrointest Surg*. 2012;16(11):2011-2025.

2. Berger MY, Olde Hartman TC, Bohnen AM. Abdominal symptoms: do they disappear after cholecystectomy? *Surg Endosc*. 2003;17(11):1723-1728.

3. Behar J, Corazziari E, Guelrud M, Hogan W, Sherman S, Toouli J. Functional gallbladder and sphincter of Oddi disorders. *Gastroenterology*. 2006;130(5):1498-1509.

4. Sherman S, Lehman GA. Sphincter of Oddi dysfunction: diagnosis and treatment. *JOP*. 2001;2(6):382-400.

5. Freeman ML, DiSario JA, Nelson DB, et al. Risk factors for post-ERCP pancreatitis: a prospective, multicenter study. *Gastrointest Endosc*. 2001;54(4):425-434.

6. Cotton PB, Garrow DA, Gallagher J, Romagnuolo J. Risk factors for complications after ERCP: a multivariate analysis of 11 497 procedures over 12 years. *Gastrointest Endosc*. 2009;70(1):80-88.

7. Cohen S, Bacon BR, Berlin JA, et al. National Institutes of Health State-of-the-Science Conference Statement: ERCP for diagnosis and therapy, January 14-16, 2002. *Gastrointest Endosc*. 2002;56(6):803-809.

8. Petersen BT. An evidence-based review of sphincter of Oddi dysfunction, I: presentations with "objective" biliary findings (types I and II). *Gastrointest Endosc*. 2004;59(4):525-534.

9. Petersen BT. Sphincter of Oddi dysfunction, II: Evidence-based review of the presentations, with "objective" pancreatic findings (types I and II) and of presumptive type III. *Gastrointest Endosc*. 2004;59(6):670-687.

10. Sgouros SN, Pereira SP. Systematic review: sphincter of Oddi dysfunction—non-invasive diagnostic methods and long-term outcome after endoscopic sphincterotomy. *Aliment Pharmacol Ther*. 2006;24(2):237-246.

11. Sherman S, Lehman G, Jamidar P, et al. Efficacy of endoscopic sphincterotomy and surgical sphincteroplasty for patients with sphincter of Oddi dysfunction (SOD); randomized controlled study. *Gastrointest Endosc*. 1994;40:125A.

12. Geenen JE, Hogan WJ, Dodds WJ, Toouli J, Venu RP. The efficacy of endoscopic sphincterotomy after cholecystectomy in patients with sphincter-of-Oddi dysfunction. *N Engl J Med*. 1989; 320(2):82-87.

13. Toouli J, Roberts-Thomson IC, Kellow J, et al. Manometry based randomised trial of endoscopic sphincterotomy for sphincter of Oddi dysfunction. *Gut*. 2000;46(1):98-102.

14. Pasricha PJ. There is no role for ERCP in unexplained abdominal pain of pancreatic or biliary origin. *Gastrointest Endosc*. 2002;56(6)(suppl): S267-S272.

15. Cotton PB. ERCP is most dangerous for people who need it least. *Gastrointest Endosc*. 2001;54(4):535-536.

16. Cotton PB, Durkalski V, Orrell KB, et al. Challenges in planning and initiating a randomized clinical study of sphincter of Oddi dysfunction. *Gastrointest Endosc*. 2010;72(5):986-991.

17. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 1998;59(suppl 20):22-33, quiz 34-57.

18. Beck AT, Steer RA, Brown GK. *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corp; 1996.

19. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand*. 1983;67(6):361-370.

20. Leserman J, Li Z, Drossman DA, Toomey TC, Nachman G, Glogau L. Impact of sexual and physical abuse dimensions on health status: development of an abuse severity measure. *Psychosom Med*. 1997; 59(2):152-160.

21. Drossman DA, Leserman J, Li Z, Keefe F, Hu YJ, Toomey TC. Effects of coping on health outcome among women with gastrointestinal disorders. *Psychosom Med*. 2000;62(3):309-317.

22. Drossman DA. The functional gastrointestinal disorders and the Rome III process. *Gastroenterology*. 2006;130(5):1377-1390.

23. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30(6):473-483.

24. Durkalski V, Stewart W, MacDougall P, et al. Measuring episodic abdominal pain and disability in suspected sphincter of Oddi dysfunction. *World J Gastroenterol*. 2010;16(35):4416-4421.

25. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics*. 1979;35(3):549-556.

26. Brawman-Mintzer O, Durkalski V, Wu Q, et al. Psychosocial characteristics and pain burden of patients with suspected sphincter of Oddi dysfunction in the EPISOD multicenter trial. *Am J Gastroenterol*. 2014;109(3):436-442.

27. Romagnuolo J, Cotton PB, Durkalski V, et al. Can patient and pain characteristics predict manometric sphincter of Oddi dysfunction in patients with clinically suspected sphincter of Oddi dysfunction? *Gastrointest Endosc*. 2014;79(5):765-772.

28. Cotton PB, Lehman G, Vennes J, et al. Endoscopic sphincterotomy complications and their management: an attempt at consensus. *Gastrointest Endosc*. 1991;37(3):383-393.

29. Freeman ML, Gill M, Overby C, Cen YY. Predictors of outcomes after biliary and pancreatic sphincterotomy for sphincter of oddi dysfunction. *J Clin Gastroenterol*. 2007;41(1):94-102.

30. Topazian M, Hong-Curtis J, Li J, Wells C. Improved predictors of outcome in postcholecystectomy pain. *J Clin Gastroenterol*. 2004;38(8):692-696.

31. Luman W, Adams WH, Nixon SN, et al. Incidence of persistent symptoms after laparoscopic cholecystectomy: a prospective study. *Gut*. 1996;39(6):863-866.

32. Black NA, Thompson E, Sanderson CF; ECHSS Group. European Collaborative Health Services Study Group. Symptoms and health status before and six weeks after open cholecystectomy: a European cohort study. *Gut*. 1994;35(9):1301-1305.

33. Vetrhus M, Berhane T, Søreide O, Søndenaa K. Pain persists in many patients five years after removal of the gallbladder: observations from two randomized controlled trials of symptomatic, noncomplicated gallstone disease and acute cholecystitis. *J Gastrointest Surg*. 2005;9(6):826-831.

34. Bar-Meir S, Halpern Z, Bardan E, Gilat T. Frquency of papillary dysfunction among cholecystectomized patients. *Hepatology*. 1984;4(2):328-330.

35. Bielefeldt K. The rising tide of cholecystectomy for biliary dyskinesia. *Aliment Pharmacol Ther*. 2013;37(1):98-106.

36. Drossman DA, Li Z, Andruzzi E, et al. US householder survey of functional gastrointestinal disorders: prevalence, sociodemography, and health impact. *Dig Dis Sci*. 1993;38(9):1569-1580.

37. Park SH, Watkins JL, Fogel EL, et al. Long-term outcome of endoscopic dual pancreatobiliary sphincterotomy in patients with manometry-documented sphincter of Oddi dysfunction and normal pancreatogram. *Gastrointest Endosc*. 2003; 57(4):483-491.

38. Price DD, Finniss DG, Benedetti F. A comprehensive review of the placebo effect: recent advances and current thought. *Annu Rev Psychol*. 2008;59:565-590.

39. Wilcox CM. Exploring the use of sham design: implications for endoscopic research. *Gastrointest Endosc*. 2008;67(1):123-127.

40. Rana JS, Mannam A, Donnell-Fink L, Gervino EV, Sellke FW, Laham RJ. Longevity of the placebo effect in the therapeutic angiogenesis and laser myocardial revascularization trials in patients with coronary heart disease. *Am J Cardiol*. 2005;95(12):1456-1459.