Gastroenterology 2018;155:752–759

CLINICAL—BILIARY

No Superiority of Stents vs Balloon Dilatation for Dominant Strictures in Patients With Primary Sclerosing Cholangitis

Cyriel Y. Ponsioen,[1] Urban Arnelo,[2] Annika Bergquist,[2] Erik A. Rauws,[1] Vemund Paulsen,[3] Paolo Cantú,[4] Ilaria Parzanese,[4] Elisabeth M. De Vries,[1] Kim N. van Munster,[1] Karouk Said,[2] Olivier Chazouillères,[5] Benoit Desaint,[5] Astrid Kemgang,[5] Martti Färkkilä,[6] Schalk Van der Merwe,[7] Werner Van Steenbergen,[7] Hanns-Ulrich Marschall,[8] Per-Ove Stotzer,[8] Douglas Thorburn,[9] Stephen P. Pereira,[9] and Lars Aabakken[3]

[1]Department of Gastroenterology & Hepatology, Amsterdam UMC, University of Amsterdam, the Netherlands; [2]Department of Gastroenterology & Hepatology, Karolinska University Hospital, Huddinge, Karolinska Institutet, Sweden; [3]Department of Gastroenterology & Hepatology, Rikshospitalet, Oslo, Norway; [4]Gastroenterology and Endoscopy Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milan, Italy; [5]Department of Hepatology, Hôpital Saint Antoine, Paris, France; [6]Department of Gastroenterology & Hepatology, Helsinki University Hospital, Helsinki, Finland; [7]Department of Gastroenterology & Hepatology, Universiteitsziekenhuis Leuven, Leuven, Belgium; [8]Department of Hepatology, Sahlgrenska University, Gothenburg, Sweden; and [9]Institute of Liver & Digestive Health, University College London and Sheila Sherlock Liver Centre, Royal Free Hospital, London, United Kingdom

BACKGROUND & AIMS: Dominant strictures occur in approximately 50% of patients with primary sclerosing cholangitis (PSC). Short-term stents have been reported to produce longer resolution of dominant strictures than single-balloon dilatation. We performed a prospective study to compare the efficacy and safety of balloon dilatation vs short-term stents in patients with non–end-stage PSC. **METHODS:** We performed an open-label trial of patients with PSC undergoing therapeutic endoscopic retrograde cholangiopancreatography (ERCP) at 9 tertiary-care centers in Europe, from July 2011 through April 2016. Patients found to have a dominant stricture during ERCP were randomly assigned to groups that underwent balloon dilatation (n = 31) or stent placement for a maximum of 2 weeks (n = 34); patients were followed for 24 months. The primary outcome was the cumulative recurrence-free patency of the primary dominant strictures. **RESULTS:** Study recruitment was terminated after a planned interim analysis because of futility and differences in treatment-related serious adverse events (SAEs) between groups. The cumulative recurrence-free rate did not differ significantly between groups (0.34 for the stent group and 0.30 for the balloon dilatation group at 24 months; $P = 1.0$). Most patients in both groups had reductions in symptoms at 3 months after the procedure. There were 17 treatment-related SAEs: post-ERCP pancreatitis in 9 patients and bacterial cholangitis in 4 patients. SAEs occurred in 15 patients in the stent group (45%) and in only 2 patients in the balloon dilatation group (6.7%) (odds ratio, 11.7; 95% confidence interval, 2.4–57.2; $P = .001$). **CONCLUSIONS:** In a multicenter randomized trial of patients with PSC and a dominant stricture, short-term stents were not superior to balloon dilatation and were associated with a significantly higher occurrence of treatment-related SAEs. Balloon dilatation should be the initial treatment of choice for dominant strictures in patients with PSC. This may be particularly relevant to patients with an intact papilla. ClinicalTrials.gov no. NCT01398917.

Keywords: Biliary; Drainage; Surgery; Temporary Stent.

Primary sclerosing cholangitis (PSC) is a chronic fibroinflammatory disease of the biliary tree of unknown origin. It is a progressive disease, which causes end-stage liver failure and is associated with an increased risk of cholangiocarcinoma. Transplant-free survival is estimated between 12 and 21 years.[1,2] Currently, therapy is limited to treatment of complications such as relieving biliary obstruction and orthotopic liver transplantation in case of end-stage liver disease.[3,4]

During the natural history of the disease, many patients experience symptoms such as pruritus, right upper quadrant pain (RUQP), fatigue, and/or bouts of fever and jaundice because of impeded biliary drainage. In approximately 60% of cases, dominant strictures (DS), which may be superimposed on diffuse ductal disease, are the principal cause of these complaints.[5] The incidence is estimated at 8%–10% annually.[6]

The 2015 guidelines from the American Society for Gastrointestinal Endoscopy (ASGE) on the role of endoscopic retrograde cholangiopancreatography (ERCP) state that patients with PSC and DS should undergo ERCP with biliary sampling to assess for presence of malignancy.[7] Moreover, benign strictures respond well to endoscopic therapy. Recommendations from a recent collaboration between the European Association for Study of the Liver and the European

Abbreviations used in this paper: **ACCS, Amsterdam Cholestatic Complaints Score; ALP, alkaline phosphatase; ASGE, American Society for Gastrointestinal Endoscopy; AST, aspartate aminotransferase; CI, confidence interval; DS, dominant stricture; DSMB, data safety monitoring board; ERCP, endoscopic retrograde cholangiopancreatography; PSC, primary sclerosing cholangitis; MRCP, magnetic resonance cholangiopancreaticography; PEP, post–endoscopic retrograde cholangiopancreatography pancreatitis; RUQP, right upper quadrant pain; SAE, serious adverse event; SF36, Short Form–36; UDCA, ursodeoxycholic acid; ULN, upper limit of normal.**

Most current article

0016-5085/$36.00
https://doi.org/10.1053/j.gastro.2018.05.034

WHAT YOU NEED TO KNOW

BACKGROUND AND CONTEXT

Dominant strictures occur in approximately 50% of primary sclerosing cholangitis patients. Since more than 20 years, there has been debate regarding the preferred modality to treat dominant strictures arising in PSC.

NEW FINDINGS

Efficacy of short-term stenting is comparable to single session balloon dilatation, but associated with a relative risk of 6.8 to develop serious adverse events compared to balloon dilatation.

LIMITATIONS

Patients underwent ERCs during follow-up only when they met the criteria of a predefined decision rule or when clinically indicated. Consequently, some recurrences may have been overlooked.

IMPACT

These findings indicate that balloon dilatation should be the initial treatment of choice for dominant strictures in PSC patients without previous sphincterotomy.

Society of Gastrointestinal Endoscopy to develop guidelines on endoscopy in PSC are in line with the ASGE statements.[8,9] The level of evidence reviewed in both guidelines is low, however, based only on retrospective series.

The best therapeutic approach to treat dominant strictures is not known. Both balloon dilatation and temporary stenting are used. From the largest retrospective series on repetitive balloon dilatation, one may infer that the recurrence-free rate at 2 years after a single session of endoscopic balloon dilatation is approximately 30%.[6]

Ponsioen et al. reported that short-term stenting with a mean duration of stent placement of 11 days is safe and effective, showing improvement of cholestatic symptoms and biochemistry in 83% of patients at 8 weeks and a re-intervention–free rate of 70% at 2 years.[5] We hypothesized that short-term stenting is superior to balloon dilatation at preventing recurrence of DS. Therefore, the aim of the present study was to compare short-term stenting vs balloon dilatation for the treatment of dominant strictures in patients with PSC with regard to cumulative recurrence-free patency, safety, and short-term improvement in cholestatic symptoms and biochemistry.

Methods

Study Design

We undertook a multicenter, open-label, randomized, 1:1 parallel group trial with a follow-up of 24 months. Eligible patients were randomized during ERCP when a DS was identified to either balloon dilatation or short-term stenting for a maximum of 2 weeks.

Participants

Eligible patients had a diagnosis of PSC according to the European Association for Study of the Liver 2009 criteria,[10] ascertained with magnetic resonance cholangiopancreaticography (MRCP), ERCP, percutaneous transhepatic cholangiography, and/or liver biopsy, were between 18 and 75 years of age, and fulfilled at least 1 of the following 5 criteria sets: (1) serum bilirubin level $>$ 3 times the upper limit of normal (ULN); (2) progression of right upper quadrant pain (RUQP), pruritus, fatigue, and/or fever attributed to acute bacterial cholangitis by at least 1 grade according to the Amsterdam Cholestatic Complaints Score (ACCS)[5] (Supplementary Table 1) within the last month, together with a 50% increase of total bilirubin and/or alkaline phosphatase (ALP) within the last 4 months and absolute value $>$ 1.2 times the ULN; (3) increase of 20% or more of total bilirubin and/or ALP within the last 4 months and absolute value $>$ 1.2 times the ULN, together with a documented dominant-appearing stricture on MRCP or ERCP $<$ 4 months before screening; (4) progression of RUQP, pruritus, fatigue, and/or fever attributed to bacterial cholangitis by at least 1 grade within the last month, together with total bilirubin and/or ALP $>$ 1.2 times the ULN and a documented dominant stricture on recent MRCP or ERCP $<$ 4 months before screening; (5) summed cholestatic complaints score of $\geq$ 3, or pruritus $\geq$ 2, or RUQP $\geq$ 2 at screening, together with total bilirubin and/or ALP $>$ 1.2 times the ULN and a documented dominant stricture on recent MRCP or ERCP $<$ 4 months before screening. Criteria sets 4 and 5 were added to the eligibility criteria to increase recruitment after additional institutional review board approval in March 2013 and May 2015, respectively. All eligibility sets were designed to reflect clinical practice with regard to indication for endoscopic intervention and to allow for detection of relevant changes during follow-up. On imaging, a dominant stricture was defined as any stricture arising in the extrahepatic or left/right main ducts that was deemed functionally relevant by the treating endoscopist/radiologist.

Exclusion criteria were prior stenting or balloon dilatation within the previous 4 months; signs of bacterial cholangitis as defined by definite cholangitis according to the criteria in Supplementary Table 2[11]; change of ursodeoxycholic acid (UDCA) therapy within 4 weeks; inability to give written informed consent; biliary cirrhosis with Child-Pugh score $\geq$ 8; estimated transplant-free survival $<$ 2 years, as calculated by a Mayo score $>$ 2; suspicion of cholangiocarcinoma, reflected by an imaging study suggestive of metastasis, MRCP with mass lesion with contrast enhancement, or rise in CA19.9 of $>$ 63 U/mL in the previous 4 months together with an absolute value $>$ 130 U/mL; signs of current malignancy other than basal cell carcinoma; life expectancy $<$ 24 months; use of antibiotics in previous 4 weeks; women pregnant at the time of screening; HIV or acute or chronic hepatitis B or hepatitis C; or substance (drug or alcohol) misuse within the previous 2 years.

Outcome

The primary endpoint was the cumulative recurrence-free rate of the primary DS(s) within 24 months in those patients who did not experience initial failure (Supplementary Figure 1). Ideally, assessment of recurrence of the treated DS(s) would require successive ERCPs at regular intervals during follow-up. Because this is not possible for obvious ethical reasons, a decision rule for repeated ERCP was determined based on either (1) recurrence of serum bilirubin level back to screening level if $>$ 3 times the ULN; (2) increase of ALP or bilirubin $\geq$ baseline level, together with increase in a cholestatic complaint $\geq$ 1

point relative to the previous visit; or (3) bouts of definitive or likely cholangitis.

To identify either serum bilirubin or ALP as an indicator for recurrence of the dominant stricture warranting repeated ERCP, an initial failure assessment at 3 months was included, (see Supplementary Material and Supplementary Figure 2). Patients who showed insufficient response with regard to improvement in cholestatic liver enzymes and/or cholestatic symptoms at 3 months would qualify as having initial failure without the requirement for repeated cholangiography.

Secondary outcomes consisted of occurrence of procedure-related complications, including serious adverse events (SAEs), as defined by the World Health Organization, and change in symptoms assessed by the ACCS, liver enzymes, and change in quality of life, as measured by the Short Form–36 (SF36), both the mental component summary and the physical component summary, at 3 months. Complications within the first 30 days after randomization were rated as not related, possibly related, or related by the treating physician. The criteria to diagnose post-ERCP pancreatitis (PEP) were according to the 1991 consensus report on PEP by Cotton et al.[12] These require hospitalization or extension of hospital admission, amylase level $>$ 3 times the ULN at 24 hours after the procedure, and abdominal pain.

Sample Size and Interim Analyses

A retrospective series from Amsterdam showed a re-intervention–free rate of 70% for short-term stenting.[5] In a previous pilot study in Amsterdam, 2-year re-intervention–free survival of 50% vs 17% for short-term stenting vs balloon dilatation was observed (unpublished data). Assuming that the re-intervention–free survival rate at 2 years would be 60% for short-term stenting vs 30% for balloon dilatation yielded a sample size of n $=$ 42 per group. Allowing for a dropout rate of 15% rendered a total sample size of n $=$ 100 to attain a power of 80% and a 2-sided α level of .05. An interim analysis by a data safety monitoring board was planned when 50% of the intended total number of study subjects had passed their 3-month visit. The DSMB was assigned to advise on early termination of the trial because of safety concerns and on futility.

Interventions

Patients were recruited and treated at 9 academic centers across Europe. Eligible patients underwent ERCP after written informed consent. Periprocedural antibiotic prophylaxis consisted of cefotaxime 1000 mg intravenously within 1 hour before the procedure and 12 and 24 hours thereafter or levofloxacin 500 mg orally twice daily within 2 hours before the ERCP and 12 and 24 hours thereafter.

At ERCP, patients were randomized only when, in the absence of purulent bile and fever $>$ 38.5°C, 1 or more dominant stricture(s) of the common bile duct, the common hepatic duct, and/or the main left or right hepatic duct were encountered and deemed amenable to both balloon dilatation or stenting by the endoscopist. Block randomization in blocks of 4, to ensure even distribution among participating centers, was performed by a Web-based electronic case record form. Follow-up time began at randomization. Before any intervention, where possible, brush cytology was obtained of any suspicious dominant stricture. When brush cytology showed cholangiocarcinoma or high-grade dysplasia, that patient was withdrawn from the study.

Sphincterotomy was performed at the discretion of the endoscopist.

Stenting

A 10F polyethylene stent was inserted and retrieved after 7 days (maximum time, 14 days) by gastroduodenoscopy. No attempt to cannulate the biliary tree was made upon retrieval, unless the brush result was suspicious for cholangiocarcinoma or high-grade dysplasia. In the case of (supra-)hilar stricturing with imminent risk of cholangitis by closing of the contralateral system, an attempt should be made to insert 2 7F stents, 1 to either system. If only 1 7F stent had been placed, a new ERCP was scheduled with the same perioperative antibiotic regimen, and an attempt was made to exchange the 7F stent for a 10F or 2 7F stents, to be retrieved after another 7 days. Balloon dilatation or Soehendra dilators (Cook Medical, Bloomington IN) to facilitate stent placement was allowed. When placement of a 10F stent or at least 2 7F stents proved impossible after 2 ERCPs, treatment was regarded as a failure.

Balloon Dilatation

DSs were dilated by placing a 4 cm 6 mm–diameter biliary dilation balloon in the stricture(s). The balloon was inflated to the point that the waist disappeared completely at fluoroscopy or to the maximum recommended by the manufacturer for 2 minutes. Longer or serial strictures must be treated with successive dilatation down to the most distal (relative) narrowing during the same session. The use of Soehendra dilators to facilitate insertion of a balloon was allowed. If balloon dilatation with a 6-mm–diameter balloon proved impossible at the first attempt, a second ERCP was scheduled after 1 week. For intrahepatic strictures, a 4-mm–diameter balloon was allowed. When balloon dilatation at the second attempt proved impossible, the treatment was regarded as an initial failure for that patient.

After the procedure, patients were kept overnight for observation and intravenous antibiotics. The next morning, vital signs were recorded, a visual analog pain scale was taken, and blood was drawn to measure hemoglobin, leucocyte, C-reactive protein, and amylase levels. In 1 center, these measurements were performed 4 hours after the procedure and, if uneventful, patients were discharged, followed by a telephone interview the next morning.

Patients were reviewed at 1 week and every 3 months thereafter for 24 months.

During each visit, symptoms were assessed by ACCS and SF36, adverse events were recorded, and routine blood samples were drawn.

When a study subject met the decision rule for secondary failure, a repeated ERCP was performed: if this showed a recurrence of the original stricture, then the primary endpoint was met. If no DS or a new DS was seen, the original treatment was still deemed successful. The new stricture could be treated at the discretion of the endoscopist, but the patient's follow-up was censored at the time of repeated ERCP. When a

Table 1. Baseline Characteristics

Characteristics	Balloon dilatation (n = 31)	Short-term stenting (n = 34)
Mean age, *y* ± SD	40 ± 14	40 ± 11
Male, n (%)	22 (71)	23 (68)
Disease duration, *y*, median (IQR)	7 (3–10)	4 (2–8)
Concurrent IBD, n (%)	25/31 (80)	25/33 (76)*
UC	23 (74)	19 (58)
Crohn	1 (3)	5 (15)
IBD-U	1 (3)	1 (3)
Previous EPT, n (%)	15 (48)	10 (29)
UDCA use, n (%)	21 (68)	26 (76)
Anti-pruritic medication, n (%)	3 (10)	5 (15)
Bilirubin,[a] *μmol/L*, median (IQR)	38 (15–108)*	24 (12–57)[d]
ALP,[b] *U/L*, median (IQR)	312 (259–470)*	302 (214–475)[e]
AST,[c] *U/L*, median (IQR)	110 (48–170)**	70 (41–111)[f]

AST, aspartate aminotransferase; EPT, endoscopic papillotomy; IBD-U, inflammatory bowel disease undefined; IQR, interquartile range; SD, standard deviation; UC, ulcerative colitis.
[a]Normal value, ≤17 μmol/L.
[b]Normal range, 40–120 U/L.
[c]Normal value, ≤40 U/L.
[d]data missing for 1 patient.
[e]data missing for 2 patients.
[f]data missing for 3 patients.

cholangiocarcinoma was discovered during follow-up, this led to study withdrawal, as did any change in UDCA therapy. Data were censored from the date of change in UDCA therapy with regard to the primary endpoint.

Data Analysis

The primary endpoint was analyzed on an intention-to-treat basis by log-rank testing. The intention-to-treat population consisted of all randomized subjects. Initial failure was compared by chi-squared testing, as was the occurrence of treatment-related adverse events, with Fisher exact test applied where appropriate. Differences in change in bilirubin, ALP, and AST levels and cholestatic symptoms at 3 months were assessed as per protocol. In the per protocol population, 1 patient in each group who did not receive treatment, 2 patients in the balloon dilatation group who dropped out early because of high-grade dysplasia in their initial brush, 1 patient in the short-term stenting group who had a change in UDCA dose 1 month after randomization, and 1 patient in the short-term stenting group who was listed for liver transplant within 3 months, were left out. Changes in biochemistry and symptoms between groups were tested by Mann-Whitney *U* test. Differences in change in SF36 score between groups at 3 months were tested with Student *t* test. Overall changes at 3 months in biochemical cholestasis and complaints were tested using the Wilcoxon paired samples signed-rank test. Potential confounders on the occurrence of treatment-related SAEs were assessed by univariate and, when appropriate, multivariate logistic regression. All data analyses were performed using SPSS, version 24 (IBM, Armonk, NY).

Ethical Issues

The trial was conducted according to the principles of the Declaration of Helsinki (Seoul 2008 version) and in accordance with the Dutch Medical Research Involving Human Subjects Act. The study protocol and amendments were approved by the institutional review boards of all participating centers. The trial was monitored by an independent monitor onsite from the Clinical Research Unit of the

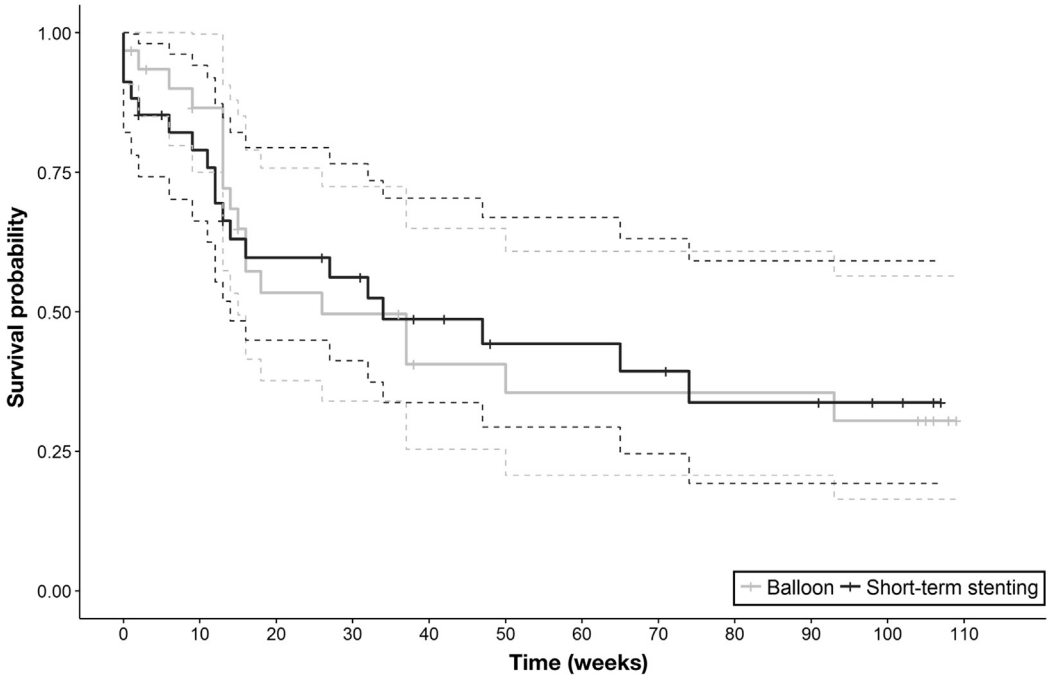


Figure 1. Cumulative recurrence-free patency of treatment. Short-term stenting, n = 34; balloon dilatation, n = 31. Log-rank test, *P* = 1.0.

Table 2. Changes in Bilirubin, ALP, and AST Levels and Cholestatic Symptoms at Baseline and 3 Months and Changes Thereof Between Groups

	Balloon Group				Stent Group				
Measure	Baseline (t_0)	3 months (t_3)	Δ ($t_0 \geq t_3$)	Relative change	Baseline (t_0)	3 months (t_3)	Δ ($t_0 \geq t_3$)	Relative change	*P* value
Bili, *μmol/L*, median (IQR)	39 (16–108) n = 30	14 (11–58) n = 28	–10 (0 to –45) n = 28	—	24 (12–57) n = 33	14 (9–24) n = 28	–3 (–1 to –33) n = 28	—	.35[a]
($Bili_0$ – $Bili_{3mo}$) / $Bili_0$, median (IQR)	—	—	—	0.30 (0–0.67) n = 28	—	—	—	0.22 (0–0.53) n = 27	.54[a]
ALP, *U/L*, median (IQR)	312 (259–470) n = 30	306 (153–446) n = 29	–31 (47 to –115) n = 29	—	302 (215–475) n = 33	280 (157–390) n = 27	–35 (21 to –133) n = 26	—	.67[a]
(ALP_0 – ALP_{3mo}) / ALP_0, median (IQR)	—	—	—	0.09 (–0.17 to 0.41) n = 29	—	—	—	0.10 (–0.12 to 0.36) n = 26	.97[a]
AST, *U/L*, median (IQR)	110 (48–171) n = 29	64 (43–175) n = 29	–12 (5 to –55) n = 28	—	70 (41–111) n = 31	44 (32–83) n = 27	–15 (–2 to –39) n = 25	—	.73[a]
(AST_0 – AST_{3mo}) / AST_0, median (IQR)	—	—	—	0.14 (–0.11 to 0.49) n = 28	—	—	—	0.25 (0.06–0.43) n = 25	.32[a]
Pruritus, median (IQR)	1.5 (0–2) n = 30	0.0 (0–1) n =	–1.0 (0 to –2) n = 28	—	1.0 (0–3) n = 33	0.5 (0–1) n = 28	0.0 (0 to –2) n = 27	—	.70[a]
Fatigue, median (IQR)	1.0 (0–2) n = 30	0.5 (0–1) n = 28	–1.0 (0 to –1) n = 28	—	1.0 (0–2) n = 33	0.0 (0–2) n = 27	0.0 (0 to –1) n = 26	—	.60[a]
Fever, n (%)	1 (3) n = 30	2 (7) n = 28	—	—	4 (12) n = 33	1 (3) n = 28	—	—	1.0[b]
RUPQ, median (IQR)	1.0 (0–1) n = 30	0.0 (0–1) n = 28	0.0 (0 to –1) n = 28	—	0.0 (0–2) n = 33	0 (0–0) n = 28	0.0 (0 to –1) n = 27	—	.68[a]
ACCS, median (IQR)	4.0 (2–4) n = 30	1.0 (0–3.5) n = 28	–1.5 (–0.25 to –3) n = 28	—	4.0 (2–5) n = 33	1.0 (0–3) n = 27	–2.0 (–1 to –3) n = 27	—	.61[a]
SF-36 MCS, mean (SD)	43.2 (10.3) n = 29	49.4 (8.5) n = 20	6.5 (11.4) n = 20	—	41.7 (11.4) n = 28	49.5 (8.3) n = 16	6.5 (8.1) n = 14	—	.99[c]
SF-36 PCS, mean (SD)	46.4 (8.5) n = 29	51.4 (5.0) n = 16	4.7 (8.3) n = 20	—	47.2 (8.7) n = 28	49.4 (8.9) n = 16	0.9 (9.4) n = 14	—	.23[c]

NOTE. Analyses were performed in the per protocol population.
AST, aspartate aminotransferase; Bili, bilirubin level; IQR, interquartile range; MCS, mental component score; PCS, physical component score; SD, standard deviation.
[a]Mann-Whitney *U* test.
[b]Fisher[1]s exact test.
[c]Independent samples T-test.

Academic Medical Center, Amsterdam. All authors had access to the study data and reviewed and approved the final manuscript.

Results

Recruitment

From July 2011 to April 2016, 80 patients were enrolled, and 65 were randomized during ERCP. Recruitment is shown in Supplementary Figure 1. One patient in each treatment arm was excluded from safety analysis because they did not receive treatment. Both were randomized before mandatory passage with a guiding catheter to show accessibility of the strictures to therapy. Also, they did not undergo a prescribed second attempt. Two patients in the balloon dilatation arm dropped out early, because their brush results proved positive for high-grade dysplasia. In April 2016, recruitment was prematurely stopped at the advice of the DSMB after a planned interim analysis because of futility and a disproportionate occurrence of treatment-related SAEs between groups. Table 1 shows demographics and clinical characteristics of the study population. Baseline data were balanced between groups. More patients in the balloon dilatation group had previously undergone a sphincterotomy, but this difference was not statistically significant ($P = .12$). Details of the findings during ERCP are shown in Supplementary Table 3. The distribution of stricture(s) found at ERCP was balanced between the 2 treatment arms.

Outcomes

Primary Endpoint. The cumulative recurrence-free rate of the primary DS(s) is depicted in Figure 1 and was equal in both groups ($P = 1.0$), with almost completely overlapping confidence intervals (CIs). Overall median recurrence-free rate was 26 weeks (95% CI, 2.3–50.0) and 34 weeks (95% CI, 4.0–64.0) for the balloon dilatation and short-term stenting groups, respectively.

Initial failure was observed in 16 of 31 (52%) and 15 of 34 (44%) patients in the balloon dilatation and short-term stenting groups, respectively ($P = .89$). Reasons for initial failure are listed in Supplementary Table 4.

Secondary Endpoints. The vast majority (40/52, 77%) of patients experienced improvement in their overall symptom scores at 3 months, which mirrored a reduction in serum bilirubin, ALP, and AST levels (Supplementary Table 5).

Table 2 shows change in ACCS, cholestatic liver enzyme levels, cholestatic symptoms, and SF36 mental component summary and physical component summary scores at 3 months compared with baseline. No differences were observed in any of the secondary endpoints between groups.

Table 3 lists the occurrence and nature of procedure-related SAEs. There was a strikingly elevated relative risk of 6.8 (95% CI, 1.7–27.4) for developing treatment-related SAEs in the short-term stenting group compared with the balloon dilatation group, predominantly driven by PEP. The univariate analysis of potential confounders with respect to PEP and cholangitis/cholecystitis is shown in Table 4. The only statistically significant association was randomization. A sensitivity analysis for PEP alone showed that in univariate analysis there was a statistically significant association of stenting with the risk for PEP as an SAE (odds ratio, 9.3; 95% CI, 1.1–79.4; $P = .04$), with an expected strong, albeit not statistically significant, protective effect of previous sphincterotomy. In multivariate analysis, carrying forward randomization and previous sphincterotomy only to avoid overfitting, we found a trend toward stenting being associated with PEP ($P = .07$), but the numbers were probably too low to show a statistically significant association (Supplementary Table 6). The risk of developing PEP requiring hospitalization in patients who had a previous EPT was only 4% (1 in the stent group), whereas this was 20% (1 in the balloon dilatation group, 7 in the stent group) in patients with intact papilla (odds ratio, 0.17; 95% CI, 0.02–1.4; $P = .10$).

Discussion

For more than 20 years, there has been debate regarding the preferred modality to treat dominant strictures arising in PSC.[5,6,8,13] To our knowledge, we present the first-ever randomized trial on endoscopic therapy in PSC. We could not confirm our hypothesis that short-term stenting is superior to balloon dilatation at preventing recurrence of DS. Indeed, although the study was not designed as a non-inferiority trial, the results showed that cumulative recurrence-free rates of DS over 2 years of follow-up after single-session balloon dilatation or short-term stenting

Table 3. Procedure-Related SAEs

Procedure	Balloon dilatation n = 30[a]	Short-term stenting n = 33[a]	OR (95% CI)	*P* value
All cause, n (%)	2 (6.7)	15 (45.4)	11.7 (2.4–57.2)	.001
Cholangitis/cholecystitis, n (%)	1 (3.3)	4 (12)	4.0 (0.42–38.0)	.36
Post-ERCP pancreatitis, n (%)	1 (3.3)	8 (24)	9.3 (1.1–79.4)	.03
Postprocedural pain, n (%)	0	2 (4.5)	n.a.	—
Ascites	0	1 (3)	n.a.	—

n.a., not applicable; OR, odds ratio.
[a]One patient in each group was left out from the comparison because they did not receive an attempt at their allotted treatment due to failure to pass the dominant stricture with a guidewire.

Table 4. Univariate Analysis for Risk Factors of Procedure-Related Serious Adverse Events

Risk factor	OR	Univariate analysis, 95% CI	*P* value
Age	1.0	0.9–1.03	.42
Sex	1.2	0.4–4.0	.71
Disease duration	1.0	0.9–1.1	.83
Randomization	11.2	2.4–57.2	.002
Center	0.8	0.6–1.1	.16
Previous sphincterotomy	0.6	0.2–2.0	.39
Baseline serum biluribin	1.0	0.99–1.01	.43
Procedure time	1.0	0.99–1.02	.79
NSAID prophylaxis	1.0	0.3–3.1	.99
Sphincterotomy	0.7	0.2–2.15	.48

NSAID, nonsteroidal anti-inflammatory drug; OR, odds ratio.

were comparable, with Kaplan-Meier curves and 95% CIs almost completely overlapping. Although only 65% of the predefined sample size was available for interim analysis, the DSMB calculated that the likelihood of yielding a significant contrast between the 2 treatment arms with 100% of the sample size included would be less than 5%.

On the other hand, there was a strong safety signal in the stenting group, who experienced almost 7 times more frequent SAEs, such as PEP, cholangitis, and cholecystitis, compared with the balloon dilatation group, whereas in the latter group the SAE rate was 6.7%, which is consistent with earlier observations.[6,14] This large contrast in occurrence of SAEs between groups was the main reason for terminating the trial prematurely. The overall rate of PEP as an SAE in this trial was 14%, which is approximately twice the rate reported in several recent retrospective series,[14,15] which used a similar definition of PEP and had comparable rates of previous sphincterotomy as the major protective factor. Putative factors explaining the difference between these series and our trial are that the former applied stenting much less to treat DSs and had a retrospective design, which may be prone to underreporting of adverse events.

The reason for the elevated risk of PEP may be the space-occupying effect of a standard 10F transpapillary stent or 2 7F stents obstructing the pancreatic duct in an often small and retracted papilla in PSC. This is supported by our and other's observations that PEP occurred much less frequently after previous sphincterotomy.[14,16,17] Likewise, 8–10-mm self-expandable metallic stents for malignant obstruction in pancreatic carcinoma or distal cholangiocarcinoma are associated with an increased risk of PEP compared with 10F plastic stents.[18] Furthermore, a stent may block the cystic duct or the contralateral main intrahepatic duct in the often multistrictured biliary tree of PSC patients, precipitating bacterial cholangitis.[16,19]

To our knowledge, our trial also yielded the first prospective data showing that both balloon dilatation and stenting of DS leads to amelioration of symptoms and significant decrease in cholestasis in the majority of patients at 3 months, strengthening the thus far low level of evidence on which the recent ASGE, European Association for the Study of the Liver, and European Society for Gastrointestinal Endoscopy recommendations for performing therapeutic ERCP in symptomatic patients are based.

An obvious advantage of balloon dilatation is that it precludes the need for further endoscopy for stent removal. However, in the largest series of balloon dilatation for DS in PSC, 1–8 (mean, 1.8 ± 0.2) repeated interventions were needed for successful opening of DSs.[6]

A limitation of the trial is that patients underwent ERCP during follow-up only when they met the criteria of a predefined decision rule or when clinically indicated, to avoid the risks of unnecessary ERCP. This generally mirrors clinical practice and justified this decision rule. Consequently, some recurrences of DSs may have been overlooked. Balloon dilatation to facilitate stent insertion was allowed and was applied in half of the patients who were randomized to receive stent placement. If the trial had yielded a significant benefit of short-term stenting compared with balloon dilatation, this could theoretically have constituted a confounder. However, because the trial was in fact comparing 2 endoscopic treatment strategies, we believed that any measure to facilitate stent placement was appropriate. Another limitation is that patients with Child-Pugh score ≥ 8 or Mayo Risk score > 2 were excluded. It is unclear if the results can be extrapolated to such more advanced cases, because these patients may respond less favorably, for example, in terms of decrease of bilirubin level. More than two thirds of patients in each group used UDCA maintenance therapy. This could potentially have blunted the biochemical response at 3 months and may have affected the relatively high proportion of patients who did not meet the initial response criteria.

In conclusion, endoscopic treatment of dominant strictures is efficacious in ameliorating symptoms in patients with PSC. Short-term stenting was not superior to balloon dilatation and was associated with a much higher occurrence of treatment-related SAEs. Balloon dilatation should be the initial treatment of choice for dominant strictures in PSC. This may be particularly relevant in patients with an intact papilla.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at https://doi.org/10.1053/j.gastro.2018.05.034.

References

1. Tischendorf JJW, Hecker H, Krüger M, et al. Characterization, outcome, and prognosis in 273 patients with primary sclerosing cholangitis: a single center study. Am J Gastroenterol 2007;102:107–114.
2. Boonstra K, Weersma RK, van Erpecum KJ, et al. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. Hepatology 2013;58:2045–2055.

3. Lazaridis KN, Larusso NF. Primary sclerosing cholangitis. N Engl J Med 2016;375:1161–1170.
4. Hirschfield GM, Karlsen TH, Lindor KD, et al. Seminar primary sclerosing cholangitis. The Lancet 2013; 382:1587–1599.
5. Ponsioen CY, Lam K, Milligen de Wit AWM, et al. Four years experience with short term stenting in primary sclerosing cholangitis. Am J Gastroenterol 1999; 94:2403–2407.
6. Gotthardt DN, Rudolph G, Klöters-Plachky P, et al. Endoscopic dilation of dominant stenoses in primary sclerosing cholangitis: outcome after long-term treatment. Gastrointest Endosc 2010;71:527–534.
7. Chathadi KV, Chandrasekhara V, Acosta RD, et al. The role of ERCP in benign diseases of the biliary tract. Gastrointest Endosc 2015;81:795–803.
8. Aabakken L, Karlsen TH, Albert J, et al. Role of endoscopy in primary sclerosing cholangitis: European Society of Gastrointestinal Endoscopy (ESGE) and European Association for the Study of the Liver (EASL) Clinical Guideline. J Hepatol 2017;66:1265–1281.
9. Aabakken L, Karlsen T, Albert J, et al. Role of endoscopy in primary sclerosing cholangitis: European Society of Gastrointestinal Endoscopy (ESGE) and European Association for the Study of the Liver (EASL) Clinical Guideline. Endoscopy 2017;49:588–608.
10. European Association for the Study of the Liver. EASL clinical practice guidelines: management of cholestatic liver diseases. J Hepatol 2009;51:237–267.
11. Bilhartz LE. Gallstone disease and its complications, Gastrointestinal and liver disease. 6th ed. Philadelphia: Saunders, 1998:958.
12. Cotton PB, Lehman G, Vennes J, et al. Endoscopic sphincterotomy complications and their management: an attempt at consensus. Gastrointest Endosc 1991; 37:383–393.
13. Aljiffry M, Renfrew PD, Walsh MJ, et al. Analytical review of diagnosis and treatment strategies for dominant bile duct strictures in patients with primary sclerosing cholangitis. HPB 2011;13:79–90.
14. Ismail S, Kylänpää L, Mustonen H, et al. Risk factors for complications of ERCP in primary sclerosing cholangitis. Endoscopy 2012;44:1133–1138.
15. Seth von E, Arnelo U, Enochsson L, et al. Primary sclerosing cholangitis increases the risk for pancreatitis after endoscopic retrograde cholangiopancreatography. Liver Int 2014;35:254–262.
16. Navaneethan U, Jegadeesan R, Nayak S, et al. ERCP-related adverse events in patients with primary sclerosing cholangitis. Gastrointest Endosc 2015;81:410–419.
17. Simmons DT. Risk of pancreatitis following endoscopically placed large-bore plastic biliary stents with and without biliary sphincterotomy for management of postoperative bile leaks. Surg Endosc 2008;22:1459–1463.
18. Coté GA, Kumar N, Ansstas M, et al. Risk of post-ERCP pancreatitis with placement of self-expandable metallic stents. Gastrointest Endosc 2010;72:748–754.
19. Gluck M, Cantone NR, Brandabur JJ, et al. A twenty-year experience with endoscopic therapy for symptomatic primary sclerosing cholangitis. J Clin Gastroenterol 2008; 42:1032–1039.

Received January 14, 2018. Accepted May 8, 2018.

Reprint requests
Address requests for reprints to: Cyriel Y. Ponsioen, MD, PhD, Department of Gastroenterology & Hepatology, Academic Medical Center, PO Box 22700, 1100 DE Amsterdam, The Netherlands. e-mail: c.y.ponsioen@amc.uva.nl; fax: +31 20 6917033.

Author contributions: Cyriel Ponsioen and Lars Aabakken designed the study; Cyriel Ponsioen, Elisabeth de Vries, and Kim van Munster analyzed the data; Cyriel Ponsioen, Erik Rauws, Urban Arnelo, Annika Bergquist, Karouk Said, Vemund Paulsen, Paulo Cantú, Ilaria Parzanese, Olivier Chazouillères, Benoit Desaint, Astrid Kemgang, Martti Färkkilä, Schalk van der Mewrwe, Werner van Steenbergen, Hanns-Ulrich Marschall, Per-Ove Stotzer, Douglas Thorburn, and Stephen Pereira recruited and treated study subjects; Cyriel Ponsioen drafted the manuscript; all authors commented on and approved the manuscript.

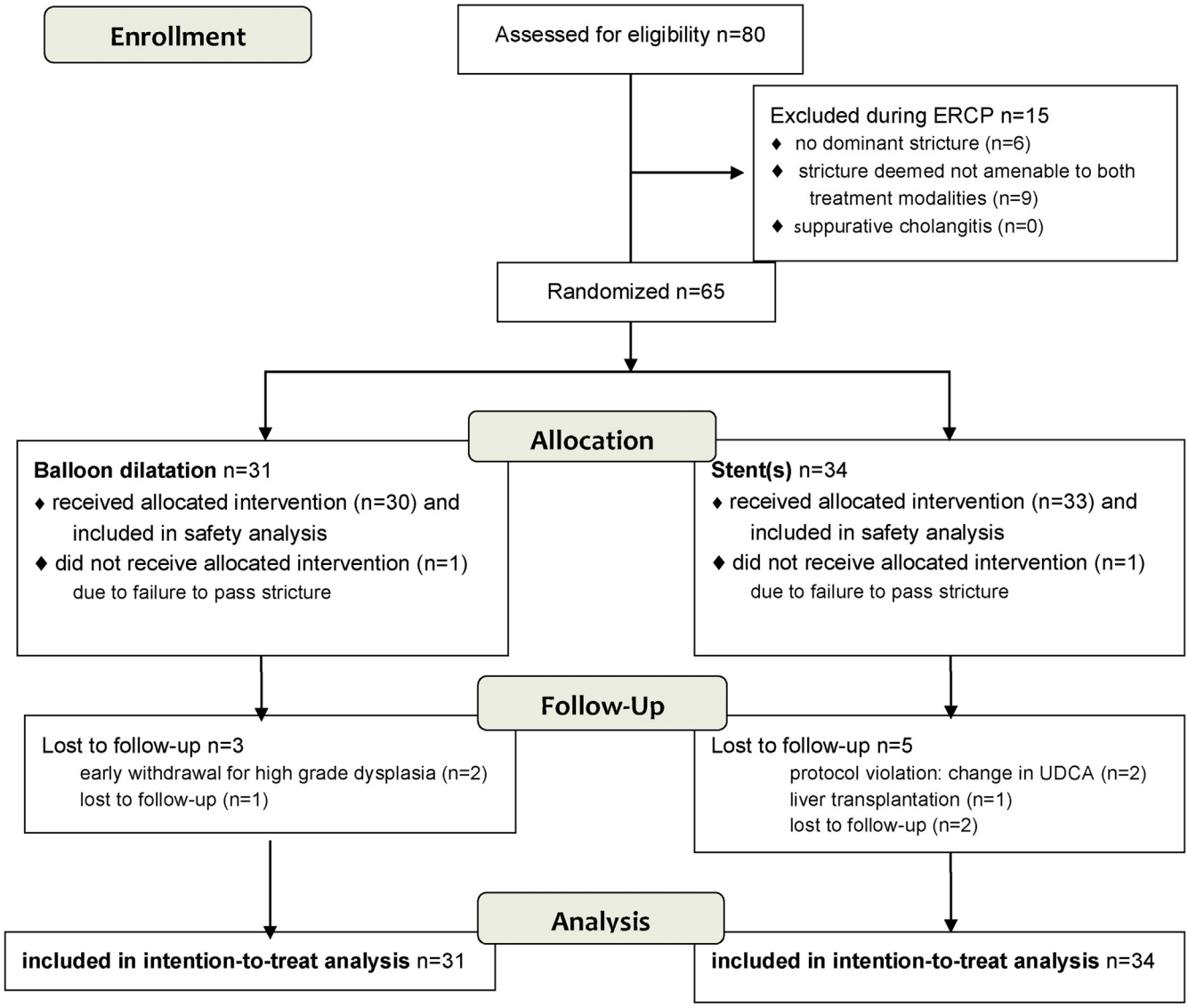
Conflicts of interest
Cyriel Ponsioen received grant support from Takeda, served in advisory boards for Takeda and AbbVie, and received speaker's fees from Takeda, AbbVie, and Dr Falk Pharma. Urban Arnelo received grant support from Boston Scientific, served on an advisory board for Boston Scientific, and received speaker's fees from Boston Scientific and Norgine Danmark A/S. Douglas Thorburn has served on advisory boards for Intercept and received speaker's fees from Intercept and Dr. Falk Pharma. The remaining authors disclose no conflicts.

Funding
This research was supported by a grant from the Norwegian PSC Foundation. The funder did not have any role in the design or conduct of the study.

Supplementary Appendix

Initial Failure Criteria

Whichever biochemical level of either ALP or bilirubin dropped more than 30% was subsequently used as the indicator for recurrence during the remainder of the follow-up. If biochemical parameters dropped by less than 30%, the treatment was designated as initial failure. This threshold of 30% was chosen on the basis of data from 1999 in the pre-UDCA era, where we found that successful treatment of a dominant stricture with a stent was commonly associated with a drop in ALP and/or bilirubin level of 30% or more.[5] However, in the trial most eligible patients used UDCA, which may be responsible for blunted dynamics in cholestatic liver enzyme levels. Planned evaluation of the first 15 patients who passed the 3-month follow-up showed that there were 3 patients who clearly improved with regard to their cholestatic complaints but dropped only 24%–28% in ALP level, so they should be designated as having initial treatment failure, despite obvious clinical improvement. To bring the initial response criteria in line with the entry criteria, it was decided to request an amendment from the institutional review board to add the following to the 30% biochemical improvement requirement: 20%–30% improvement together with improvement of at least 1 point in ACCS (Supplementary Figure 2). This became effective after institutional review board approval in May 2013 and had no bearing on the aforementioned 3 study subjects, because they were still in follow-up. Other reasons for initial treatment failure were procedure-related complications necessitating early re-intervention, assigned treatment failure, and recurrence or persistence of the treated stricture upon early repeated ERCP.



Supplementery Figure 1. DILSTENT flow diagram.

Supplementary Table 1. Amsterdam Cholestatic Complaints Score[5]

	0	1	2	3	4
Pruritus	None	Sometimes	Daily	Wakes me up/need medication	Intolerable
Fatigue	None	Cannot do everything	Have to rest	<50% daytime in bed	Fully bedridden
Cholangitis	No fever	Fever			
RUQ pain	None	Sometimes	Daily	Wakes me up/need analgesics	Intolerable

RUQ, right upper quadrant.

Supplementary Table 2. Criteria for Bacterial Cholangitis According to Bilhartz[11]

Major criteria
- Body temperature $> 38^{\circ}C$
- At least 1 positive blood culture result
- Clinical signs of sepsis syndrome (chills, hypotension [systolic blood pressure $<$ 90 mm Hg, diastolic blood pressure $<$ 60 mm Hg] and tachycardia [heart rate >100 beats per minute])

Minor criteria
- A rise in cholestatic enzyme levels (elevation in total bilirubin, alkaline phosphatase, or γ-glucuronyl transferase level by more than 1.5 times compared with pre-ERCP values or upsloping of these values after initial improvement above 1.5 times of the lowest value measured)
- No symptoms pointing to infection outside the biliary tree
- Pain in the right upper quadrant of the abdomen
- Leucocytosis (defined as $>12 \times 10^9$ leukocytes/L)

Definition of cholangitis
- Definite cholangitis: 3 major or at least 2 major + 2 minor criteria
- Likely cholangitis: 2 major + 1 minor criteria
- Possible cholangitis: 1 major + at least 1 minor criteria or 2 major criteria without any minor criteria

Cholangitis unlikely: all other

Supplementary Table 3. Findings During ERCP

Finding	Balloon dilatation n = 31	Short-term stenting n = 34	*P* value
Duration of intervention, *min*, mean ± SD	58 ± 34 (n = 28)	69 ± 36 (n = 34)	.20
Second attempt needed, n (%)	2 (6) (n = 31)	3 (9) (n = 34)	.92
Sphincterotomy, n (%)	12 (39) (n = 31)	12 (35) (n = 34)	.78
Periprocedural NSAID, n (%)	18 (58) (n = 31)	21 (62) (n = 34)	.76
Distribution of stricture(s), n (%)			
CBD	16 (52) (n = 31)	23 (68) (n = 34)	.72
CHD	15 (48) (n = 31)	17 (50) (n = 34)	
Hilar	13 (45) (n = 29)	13 (38) (n = 34)	
LHD	10 (34) (n = 29)	13 (38) (n = 34)	
RHD	11 (38) (n = 29)	9 (26) (n = 32)	
Brush result, n	(n = 28)	(n = 27)	
Not suspicious	26	27	
HGD or CCA	2	0	
Soehendra dilatation, n (%)	4 (13) (n = 31)	4 (12) (n = 34)	>.99
Balloon dilatation before stent insertion, n (%)	—	18 (53) (n = 33)	
Stents, n (%):		(n = 34)	
1 × 10F		26 (76)	
2 × 10F		1 (3)	
1 × 7F		1 (3)	
2 × 7F		4 (12)	
Per-procedural complications, n	(n = 31)	(n = 34)	
Significant bleeding	0	0	
Perforation	0	0	
False route	1	1	
Second attempt needed, n (%)	2 (6) (n = 31)	4 (12) (n = 34)	.67

NOTE. n indicates numbers of patients for whom each particular value was known.
CBD, common bile duct; CCA, cholangiocarcinoma; CHD, common hepatic duct; HGD, high grade dysplasia; LHD, left hepatic duct; NSAID, nonsteriodal anti-inflammatory drug; RHD, right hepatic duct; SD, standard deviation.

Supplementary Table 4. Reasons for Initial Failure

Reason	Balloon dilatation n = 29[a]	Short-term stenting n = 32[b]
Total bilirubin or alkaline phosphatase level insufficiently decreased	10	7
Assigned treatment failure	1	2
Procedure-related complications	0	2
Recurrent/persistent DS	3	2

[a]Two additional patients dropped out early because of HGD finding in the brush results.
[b]One additional patient was referred for liver transplantation within 3 months, and 1 additional patient was censored early because his UDCA dose was reduced 1 month after randomization.

Supplementary Table 5. Changes in Bilirubin, ALP, and AST Level and Cholestatic Symptoms at Baseline, 3 Months, and Changes Thereof in All Patients as per Protocol

	t = 0 mo	t = 3 mo	Ratio (0–3 mo)/0 mo	Δ (0 – >3 mo), paired samples	*P* value[a]
Bilirubin, *μmol/L*, median (IQR)	28 (13–79)	14 (11–26)	0.29 (0–0.55)	–5 (0 to –40)	<.001
	n = 63	n = 56	n = 55	n = 55	
ALP, *U/L*, median (IQR)	312 (227–471)	293 (158–444)	0.09 (–0.15 to 0.39)	–31 (–35 to –126)	.013
	n = 63	n = 56	n = 55	n = 55	
AST, *U/L*, median (IQR)	99 (45–146)	56 (35–133)	0.16 (0–45)	–15 (0 to –50)	<.001
	n = 60	n = 56	n = 53	n = 53	
Pruritus, median (IQR)	(0–3)	0 (0–1)		–1.0 (0 to –2)	<.001
	n = 63	n = 56		n = 55	
Fatigue, median (IQR)	1.0 (0–2)	0.0 (0–2)		0.0 (0 to –1)	.004
	n = 63	n = 55		n = 54	
RUQP, median (IQR)	0.0 (0–1)	0.0 (0–0.75)		1.0 (0 to –1)	.01
	n = 63	n = 56		n = 55	
ACCS total, median (IQR)	4.0 (2–5)	2.0 (1–3)		–2.0 (–1 to –3)	.001
	n = 63	n = 55		n = 55	
SF-36 MCS score, mean (SD)	47.0 (8.1)	50.1 (6.9)		3.1	.047[b]
	n = 57	n = 36		n = 34	
SF-36 PCS score, mean (SD)	43.2 (9.7)	49.7 (7.8)		6.5	.001[b]
	n = 57	n = 34		n = 34	

NOTE. n denotes the number for which the pertaining parameter was available.
AST, aspartate aminotransferase; IQR, interquartile range; MCS, mental component score; PCS, physical component score; SD, standard deviation.
[a]Wilcoxon Rank test.
[b]Paired samples *t* test.

Supplementary Table 6. Univariate and Multivariate Analysis of Risk Factors for Post-ERCP Pancreatitis

Risk Factors	Univariate analysis, OR	95% CI	*P* value	Multivariate analysis, OR	95% CI	*P* value
Age	1.0	0.92–1.04	.45	—	—	—
Sex	1.1	0.24–4.90	.91	—	—	—
Disease duration	0.8	0.62–1.02	.07	—	—	—
Randomization	9.3	1.1–79.4	.04	7.5	0.9–66.1	.07
Center	0.9	0.6–1.3	.65	—	—	—
Previous sphincterotomy	0.2	0.02–1.40	.10	0.2	0.03–21.00	.19
Baseline serum biluribin	1.0	0.99–1.01	.44	—	—	—
Procedure time	1.0	0.98–1.02	.80	—	—	—
NSAID	0.7	0.15–3.00	.60	—	—	—
Sphincterotomy	0.8	0.2–3.8	.80	—	—	—

NSAID, nonsteroidal anti-inflammatory drug; OR, odds ratio.