



# MONASH University

## **ERCP, sphincterotomy and immediate vs delayed balloon sphincteroplasty for extraction of large biliary stones – a randomised controlled trial**

*Andrew Ginn-Ming Hardley*

*MBBS (Hons), BMedSci, FRACS*

A thesis submitted for the degree of Master of Surgery at  
Monash University in 2018  
Eastern Health Clinical School

## **Copyright notice**

© The author (2018).

*I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.*

## Abstract

Endoscopic Retrograde Cholangiopancreatography (ERCP) with endoscopic sphincterotomy is the most common method for extracting biliary stones. Large biliary stones are challenging to extract endoscopically, with around 15% being too large to allow extraction after sphincterotomy alone. The addition of large-balloon papillary dilatation (sphincteroplasty) after sphincterotomy was developed to further enlarge the biliary orifice, facilitating extraction of such large stones. When this dilatation was first described as being performed at the same session as sphincterotomy, this was received with some trepidation, given that a marked increase in complications had been previously demonstrated when performing sphincteroplasty alone as an initial procedure. This concern led sphincteroplasty to be more commonly delayed to a second, subsequent ERCP. As this technique was examined and demonstrated to be safe and effective, the question arose: is it more efficacious and just as safe to perform these two components together, or are we better to delay dilatation to a second, staged procedure?

**Methods:** Between June 2015 & September 2016, we performed a non-blinded randomised controlled trial. We enrolled 150 patients undergoing primary ERCP at two tertiary-hospital surgeon-led endoscopy units, where the indication for ERCP was (potentially large) biliary stones. 52 patients with large biliary stones ( $\geq 8\text{mm}$ ) were eligible and randomised to either *a*) combined sphincterotomy + immediate balloon sphincteroplasty (immediate dilatation arm) or to *b*) our previous standard practice of sphincterotomy and stenting at the initial ERCP, with the balloon sphincteroplasty at a subsequent procedure (control arm). Demographic data, indications and procedural characteristics were compared & identical between arms. Efficacy and safety (complications) were compared between arms across a number of measures.

**Results:** We found performing a combined procedure was more efficacious in stone extraction (decreased number of procedures to clear duct  $p < 0.001$ ), with no increase in overall complication rate (15% both arms) or Comprehensive Complication Index (mean 3.8 vs 3.4,  $p = 0.87$ ). In addition, use of a combined approach lead to significant improvements in overall procedural duration (52.0 vs 70.3 mins,  $p = 0.03$ ) and total radiation dose (9.5 vs 17.6mGy,  $p = 0.03$ ).



**Conclusion:** for ERCP extraction of large biliary stones, combining sphincterotomy & balloon sphincteroplasty in a single procedure is more efficacious and as safe as performing these sequentially using a staged, two-procedure approach.

## **Declaration**

---

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signature:

Print Name: Andrew Hardley

Date: 30/11/18

## Table of Contents

<b>Copyright notice</b>	<b>3</b>
<b>Abstract</b>	<b>4</b>
<b>Table of Contents</b>	<b>7</b>
<b>Acknowledgements</b>	<b>9</b>
<b>Introduction</b>	<b>12</b>
<i>Context</i>	12
<i>Gap in knowledge</i>	12
<i>Brief methodology</i>	12
<i>Outline of chapters</i>	13
<b>Literature review</b>	<b>15</b>
<i>Gallstone disease</i>	15
<i>ERCP</i>	15
<i>Complications of ERCP</i>	17
<i>Balloon sphincteroplasty vs sphincterotomy – the techniques</i>	17
<i>Limits of sphincterotomy</i>	21
<i>Sphincterotomy + large balloon sphincteroplasty – the combined technique</i>	24
<i>Immediate vs delayed biliary dilatation</i>	32
<i>Efficacy measurement</i>	33
<i>Complication measurement</i>	35
<b>Methodology</b>	<b>38</b>
<i>Hypothesis</i>	38
<i>Sample size calculation</i>	38
<i>RCT design</i>	39
<i>Randomisation procedure</i>	42
<i>Arms</i>	43
<i>Antibiotic and NSAID protocol</i>	44
<i>Measurement of Complications:</i>	44

<i>Data collection</i>	45
<i>Analysis techniques</i>	46
<b>Results</b>	<b>47</b>
<i>Enrolment</i>	47
<i>Demographics</i>	49
<i>Inclusion criteria</i>	53
<i>Procedural details</i>	54
<i>Efficacy</i>	58
<i>Time</i>	59
<i>Safety</i>	61
<b>Discussion</b>	<b>67</b>
<i>Introduction</i>	67
<i>Procedural performance</i>	67
<i>Efficacy</i>	68
<i>Secondary endpoints of efficacy</i>	74
<i>Safety</i>	77
<b>Conclusions</b>	<b>84</b>
<i>Efficacy</i>	84
<i>Safety</i>	85
<i>Subpopulations that may not be appropriate/advantageous</i>	85
<i>Limitations</i>	86
<i>Strengths</i>	87
<i>Future directions</i>	89
<b>Appendices</b>	<b>91</b>
<i>Appendix A: Eastern Health Antibiotic and NSAID protocol for ERCP prophylaxis</i>	91
<i>Appendix B: Cotton criteria</i>	92
<i>Appendix C: Clavien-Dindo score</i>	93
<b>Bibliography</b>	<b>97</b>

## Table of Figures

Figure 1: ERCP .....	16
Figure 2: the sphincterotome sitting within the ampulla. ....	18
Figure 3: Needleknife papillotomy. ....	19
Figure 4: Sphincteroplasty image.....	20
Figure 5: Flat papilla .....	22
Figure 6: Large papilla.....	22
Figure 7: Step 1: Selective deep biliary cannulation of the bile duct in the standard fashion	24
Figure 8: Step 2: Partial sphincterotomy .....	25
Figure 9: Step 3: Balloon dilatation of the entire distal duct.....	25
Figure 10: Image intensifier image of dilating balloon. ....	26
Figure 11: Key sphincteroplasty papers.....	27
Figure 12: Cotton Criteria (adapted) .....	35
Figure 13: Estimated proportional success for each arm .....	38
Figure 14: Inclusion criteria diagram.....	41
Figure 15: CONSORT diagram.....	48
Figure 16: Demographic Data .....	49
Figure 17: Smoking status .....	50
Figure 18: ASA (American Society of Anaesthesiologists) score .....	51
Figure 19: ASA distribution .....	51
Figure 20: Indication for ERCP .....	52
Figure 21: Sphincterotomy Technique .....	56
Figure 22: Procedural data.....	57
Figure 23: Number of ERCPs performed.....	58
Figure 24: Complication Summary .....	61
Figure 25: Radiation dose.....	64
Figure 26: Radiation dose histogram.....	65
Figure 27: Radiation dose .....	66
Figure 28: Total Number of ERCPs .....	70

Figure 29: Stone caught in junction between cystic and bile ducts.....	72
--	----

## **Acknowledgements**

I would like to express my gratitude to the following people for their assistance and support in the running of this trial and preparation of the thesis.

### ***Supervisors***

Dr Sean Mackay

Prof Ian Davis

First and foremost I would like to thank my supervisors for their continual advice and support, assistance with statistical analysis and patient guidance through the preparation of this thesis.

### ***Endoscopists***

Mr Simon Banting

A/Prof Richard Cade

Mr Adrian Fox

Mr Sayed Hassen

Dr Sean Mackay

For their support of the concept and running of the trial and performance of the required ERCP procedures. In addition, they collaborated in designing inclusion/exclusion criteria and the procedural protocols for the 2 arms.

### ***Hepatobiliary/Upper GI Surgical Fellows of Box Hill and St Vincent's Hospitals***

Mr Ben Keong

Ms Sheryn Cheah

Mr Lawrence Lau

For assistance with participant recruitment, performance of required ERCP procedures and documentation of procedural details, as well as perioperative care of the participants.

### ***Endoscopy unit staff of both Box Hill and St Vincent's Hospitals, Melbourne.***

For their support of the performance of the trial, care of the patients and assistance with documentation of procedural details.

## **Introduction**

### **Context**

Bile duct stones are a common problem, which can cause symptoms ranging from mild discomfort and deranged liver function, to life-threatening pancreatitis and biliary sepsis. Bile duct stones are most commonly removed endoscopically by ERCP. To allow stone extraction, the distal end of the bile duct needs to be widened, most commonly by cutting the muscle of the sphincter at the distal end of the bile duct (sphincterotomy).

There is a limit to how wide a sphincterotomy can be safely cut and around 15% of stones are too large to be extracted by sphincterotomy alone. Traditionally these have required major open surgery to explore the bile duct and extract the stones. A new technique has been developed in which a dilating balloon is used to further widen the biliary orifice (sphincteroplasty) in patients who have already had a sphincterotomy. This combination sphincterotomy + sphincteroplasty approach was introduced with some trepidation due to previous experience with sphincteroplasty as a stand-alone procedure, which was associated with a markedly increased risk of causing pancreatitis.

### **Gap in knowledge**

A substantial number of studies have described performing sphincteroplasty after sphincterotomy. Whilst these studies are limited by their predominantly retrospective nature, this technique appears to allow extraction of larger stones than sphincterotomy alone and appears to be similarly safe. The sphincteroplasty component may be performed either immediately in the same procedure as the sphincterotomy or delayed to a second procedure, as was our current practice at the time of the trial. It is unclear which of these approaches is most appropriate and this decision is dependent on the balance between their relative efficacies and safety. Current evidence to guide our approach is limited to only two non-randomised retrospective case series.

### **Brief methodology**

We performed a non-blinded, randomised controlled trial, aimed at comparing sphincterotomy and either immediate or delayed sphincteroplasty. Participants with large bile duct stones undergoing their first ERCP were randomised in a 1:1 ratio to one of two options: a) sphincterotomy, followed by immediate sphincteroplasty and attempted stone extraction or b) sphincterotomy, followed by stent placement and return at a second ERCP



procedure for delayed sphincteroplasty and attempted stone extraction – our previous standard practice. In essence, the two arms receive the same interventions, but the two components are separated by time in the control arm (b).

We compared the two arms over a number of measures of efficacy, primarily the number of ERCPs required for duct clearance, but also the proportion of patients in whom the duct is cleared at first attempt, total duct clearance and a number of secondary markers of efficiency of resource utilisation: length of stay, total procedural time and anaesthetic time.

These outcomes with regards to efficacy were balanced against comparisons of safety. The overall complication rates were compared between arms, as well as those of each of the predictable complications of ERCP (pancreatitis, cholangitis, bleeding and perforation) and the risks of ERCP-related radiation to both patient and endoscopy staff alike. In addition, a novel metric, the Comprehensive Complication Index is used to provide a more nuanced way of comparing the disparate complications of ERCP.

## **Outline of chapters**

### **Literature review**

The thesis begins by exploring the conceptual basis for the procedure of ERCP, sphincterotomy and sphincteroplasty, and the variations used in the arms of the trial. In particular, we explore the evolution of ERCP for stone extraction, the merits of sphincterotomy and sphincteroplasty techniques alone and concerns regarding excess complications (particularly the increased incidence of pancreatitis) with the technique of sphincteroplasty. The development of the combined technique of sphincterotomy + sphincteroplasty is described, along with the evidence base for its use and the limitations therein. Finally, the theoretical concerns of combining both sphincterotomy and dilatation into a single procedure are discussed.

### **Methodology**

This chapter extrapolates upon the inclusion and exclusion criteria, randomisation protocol and procedural details of each arm, as well as procedures for data collection and techniques of data analysis.

## **Results**

Here our results are described, providing comparison between arms based on demographic information, technical details of each procedure as well as comorbidities and indications for ERCP, demonstrating equivalence between the two arms.

The arms are compared regarding efficacy over multiple endpoints: in particular, the primary endpoint of number of ERCPs required to clear the bile duct, but also procedural and anaesthetic time and length of stay. Efficacy is contrasted with the complications of each arm. This incorporates overall complication rate, descriptions of individual complications, radiation exposure and the Comprehensive Complication Index.

## **Discussion**

This section begins by benchmarking our performance of ERCP overall against appropriate large series, demonstrating appropriate technical proficiency within our practice. It then discusses the rationale behind our choice of primary endpoint in measurement of efficacy and compares the two arms across all measures of efficacy used. The various possible measures of efficacy are discussed, using multiple measures of efficacy to benchmark our trial against the literature and discussing the practical importance of each alternate measure of efficacy.

The clinical significance of our results for the various measures of efficacy, efficiency of resource utilisation and safety are discussed, identifying significant differences between approaches. The limitations of our data, in particular in comparing complications, are explored, along with the use of the Comprehensive Complication Index in an attempt to improve on these.

## **Conclusions**

To conclude, the trial is summarised, discussing the strengths and limitations of our study. Finally, the thesis identifies directions for future investigation including the role of prospectively gathered large-scale audit data.

## **Literature review**

### **Gallstone disease**

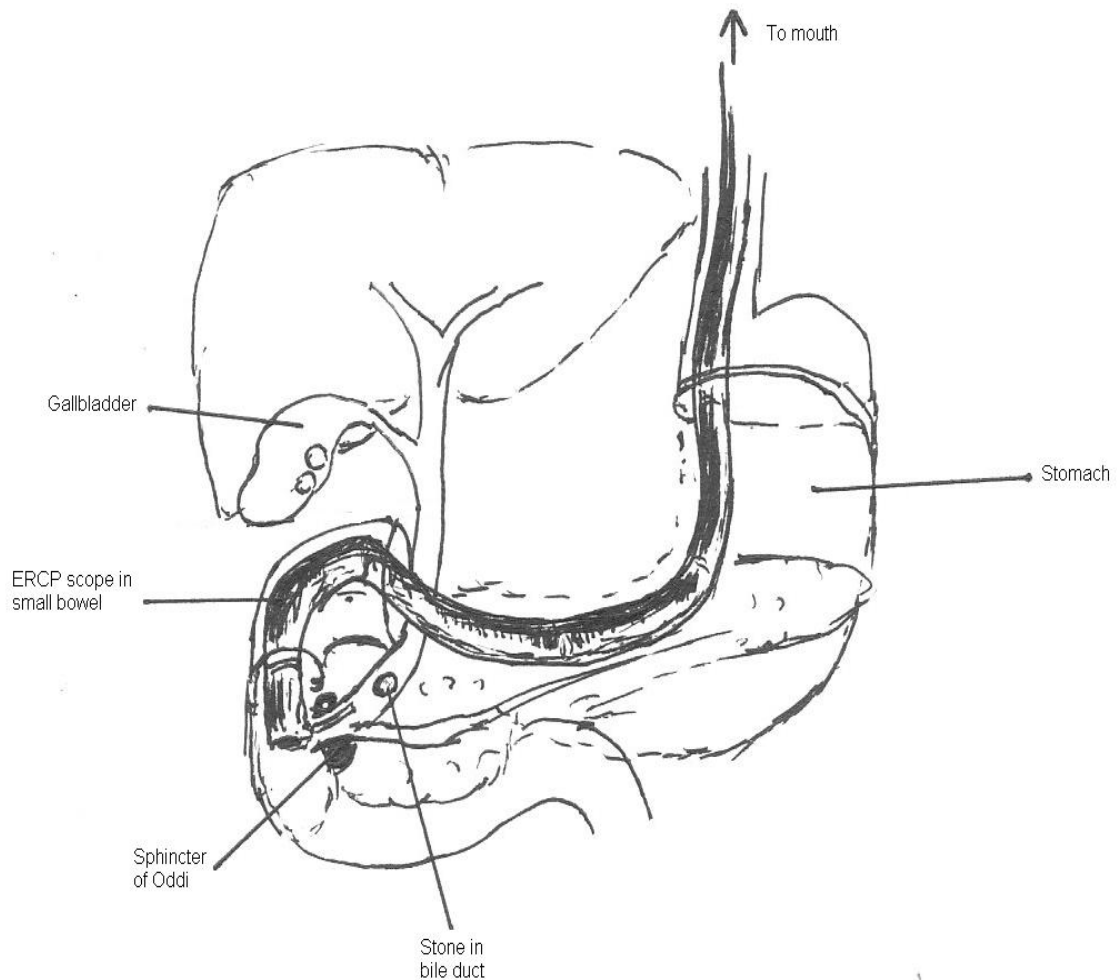
Gallstones are stones formed anywhere within the biliary system, but most commonly within the gallbladder. Gallstones are common and estimated to be present in around 15% of the population. They are often asymptomatic, but may cause significant complications, making removal of the gallbladder (cholecystectomy) one of the most common general surgical operations(1). The clinical problems caused by gallstones vary, depending on their position.

Stones in the bile ducts (choledocholithiasis) may cause significant potential morbidity. These may cause obstructive jaundice, gallstone pancreatitis and infection of the bile ducts (cholangitis), which may be life-threatening. Removal of bile duct stones may be performed either at open surgery, laparoscopically, or most commonly endoscopically.

### **ERCP**

Endoscopic retrograde cholangiopancreatography (ERCP) was initially introduced as a diagnostic procedure. It involves passing a side-viewing endoscope (duodenoscope) into the duodenum, then passing a catheter into the common bile duct (a process known as cannulation) (see Figure 1). This catheter can inject water-soluble contrast directly into the bile duct to allow x-ray imaging of the biliary tree. Likewise, instruments can be passed into the bile duct, allowing manipulation and extraction of bile duct stones, amongst other therapeutic techniques.

**Figure 1: ERCP**



New techniques have evolved to image the bile duct in a non-invasive fashion. Imaging techniques, such as Magnetic Resonance Cholangiopancreatography (MRCP) have largely replaced ERCP as a pure diagnostic tool. ERCP has evolved to become primarily a therapeutic tool and is now by far the most common method of extracting biliary duct stones.

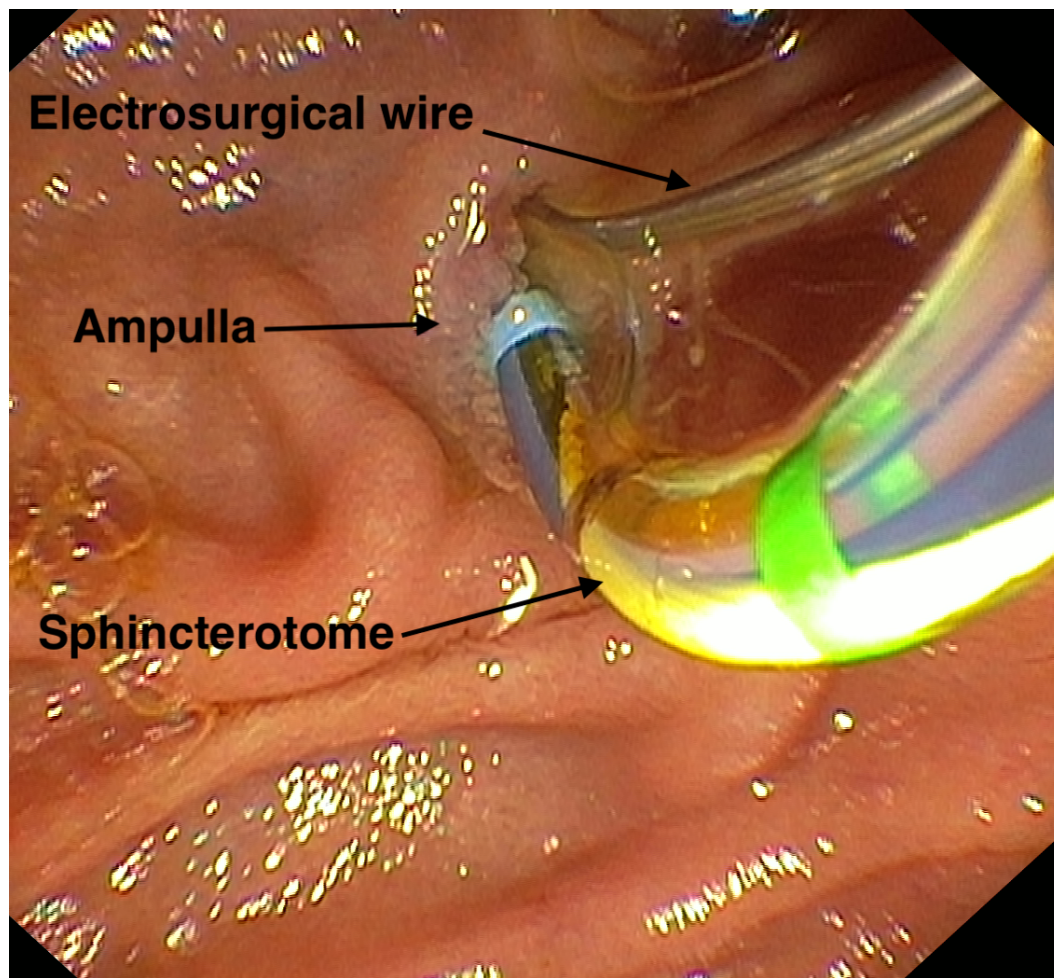
## **Complications of ERCP**

ERCP is perhaps the most challenging and dangerous endoscopic modality, with complications occurring in around 7% of patients(2). Manipulation of the ampulla of Vater can lead to mechanical pancreatic irritation, as well as reflux of biliary and duodenal content into the pancreatic duct, leading to pancreatitis. Cholangitis (bile duct infection) can occur from introduction of enteric bacteria into the usually sterile biliary tree, especially if obstruction to biliary drainage is present. Bleeding may occur, particularly from therapeutic intervention at the ampulla and can occur up to 10 days post-procedure. Finally and most concerning, there is a small but life-threatening risk of perforation, either of duodenum or oesophagus. This may occur either from the mechanical passage of the duodenoscope itself, by passage of the wire through the duodenal wall during cannulation attempts or from full-thickness division of the duodenal wall during attempts to widen the ampullary orifice(3).

## **Balloon sphincteroplasty vs sphincterotomy – the techniques**

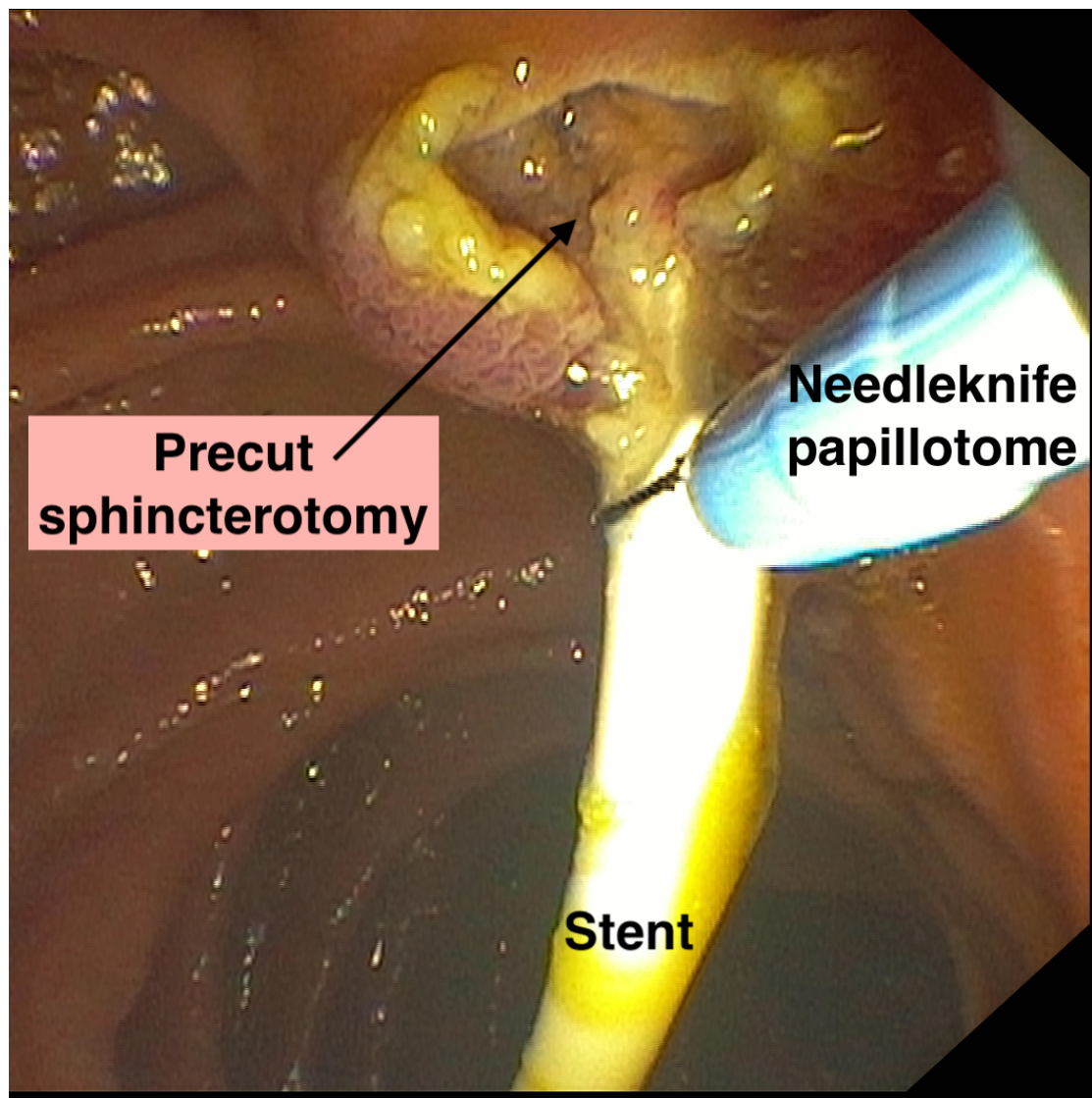
### **Sphincterotomy**

In order to extract stones, the biliary orifice, the papilla or ampulla of Vater must be enlarged. First described in 1974(4, 5), endoscopic biliary sphincterotomy is the most commonly used means of enlarging the biliary orifice to allow stone retrieval. It involves selective division of part of, or the entire biliary portion of the ampullary sphincter complex. This generally involves deep cannulation of the bile duct, wire placement under x-ray control to confirm biliary cannulation and guide further intervention, then division of the biliary sphincter with electrocautery using a specifically designed tool – the papillotome, also known as the sphincterotome. This is a catheter with an electrosurgical wire at the distal end (see Figure 2) that allows both flexion of the catheter to assist cannulation and cutting by passing a current through the wire.



**Figure 2: the sphincterotome sitting within the ampulla. The electrosurgical wire is positioned against the sphincter, facilitating its division (sphincterotomy)**

If the bile duct is unable to be cannulated initially, a sphincterotomy can be cut freehand (precut), either with the standard papillotome or with a different tool – the needle-knife papillotome (see Figure 3). Cutting without the guidance of a wire within the bile duct is more difficult and may be associated with higher risks of complications(3). Nevertheless this may be required if one is unable to place a guide-wire into the bile duct to guide sphincterotomy in the standard fashion. Estimates vary from centre to centre as to how often this is required, but according to Sweden's comprehensive population-based registry of over 11,000 ERCPs, approximately 12% of sphincterotomies require a precut technique(6).



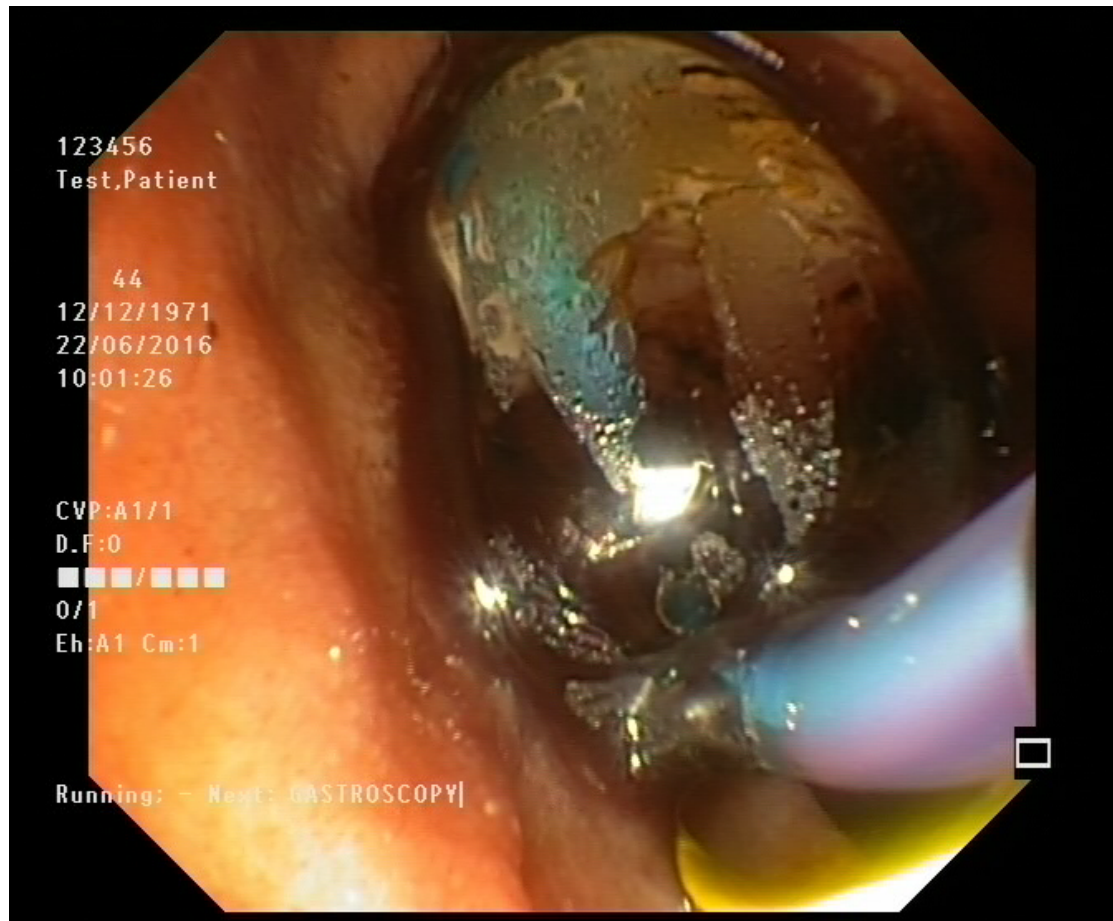
**Figure 3: Needleknife papillotomy.** In this case a sphincterotomy has been cut freehand (precut) over a stent

Complications of sphincterotomy account for a significant portion of the morbidity of ERCP. In particular, the major potential complications include pancreatitis, bleeding and duodenal perforation are all significantly more common with sphincterotomy. These all increase if it is necessary to perform sphincterotomy in a precut or needle-knife fashion.



## Sphincteroplasty

The alternate method to enlarge the biliary orifice is balloon sphincteroplasty. First described in 1982(7), this technique involves deep cannulation of the bile duct, wire placement within the bile duct and then dilatation of the biliary sphincter with a balloon passed over or alongside the guide-wire. As there is no cutting, this method theoretically causes less bleeding, a theory borne out in large meta-analyses(8).



**Figure 4: Sphincteroplasty.** Here the dilating balloon is seen passed alongside a guide-wire into the bile duct

Concern was raised about other potential complications with this technique. This was seen most famously in DiSario's randomised controlled trial comparing sphincteroplasty to sphincterotomy across 237 participants. The study was terminated at the first interim analysis due to a marked increase in pancreatitis in the sphincteroplasty arm (15.4% compared with 0.8%,  $p < 0.001$ ). There were two deaths due to severe pancreatitis in the sphincteroplasty group(9).

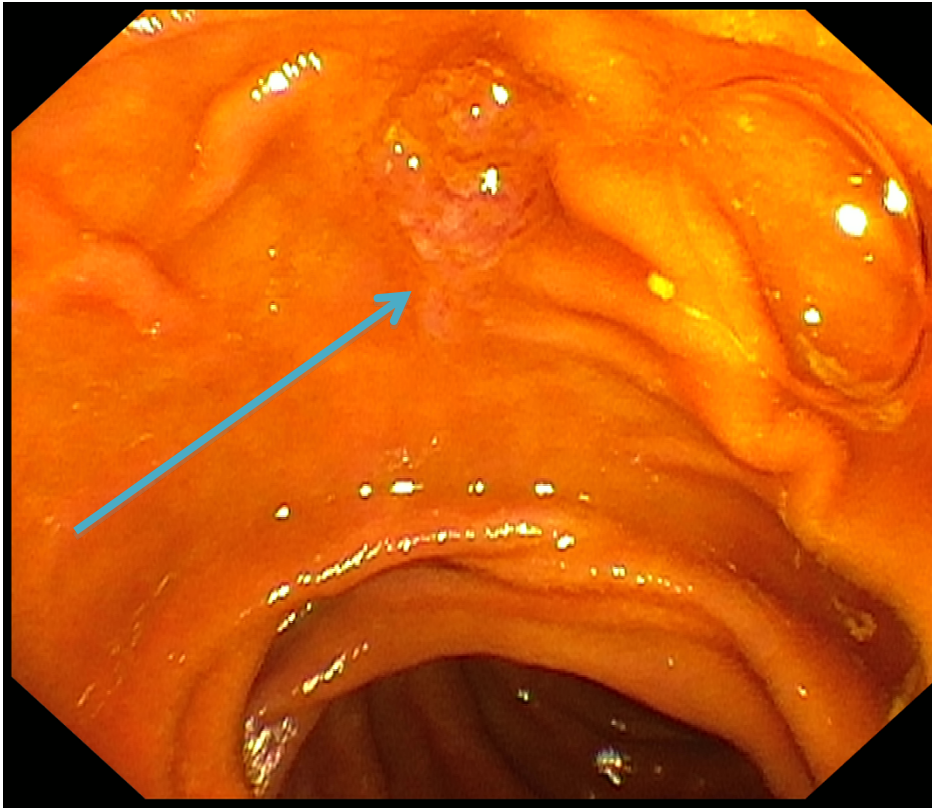


The concern regarding pancreatitis was confirmed by Weinberg's meta-analysis of 1768 patients across 15 randomised-controlled trials. This shows balloon sphincteroplasty to be associated with a two-fold increased risk of pancreatitis and slightly lower extraction rates when compared to sphincterotomy(10). This is presumed due to the mechanical compression of the pancreatic duct and sphincter by the balloon itself(11). This increase in severe pancreatitis in particular has led to sphincteroplasty as a stand-alone procedure being largely abandoned, meaning that sphincterotomy has become the dominant means of enlarging the biliary orifice to allow stone extraction.

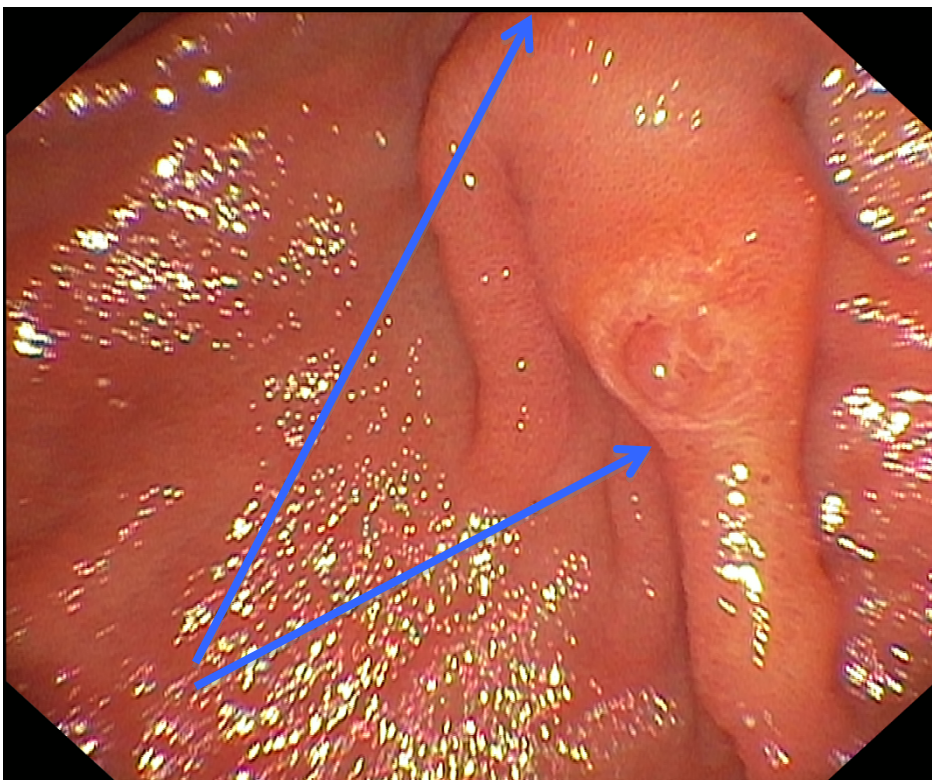
### **Limits of sphincterotomy**

Despite sphincterotomy, some stones are unable to be retrieved. A safe sphincterotomy is limited to division of the portion of biliary sphincter muscle within the duodenum. Further cutting will divide the wall of the duodenum itself, leading to perforation. The length of cut that is safely achievable will vary between individuals, as the ampulla is not uniform across the population.

Predominantly, stones that are unable to be retrieved by sphincterotomy are large – 8 – 10mm or greater. In addition, the intra-duodenal portion of the sphincter may be small (see Figure 5) or within a diverticulum, where the wall of the duodenum is particularly thin, leading to increased risk of perforation if wide sphincterotomy is performed. In these cases, even smaller stones may be impossible to retrieve by sphincterotomy alone. Other factors relating to failure of stone extraction include multiple stones, barrel-shaped stones or stones stuck above a tapering or tortuous distal bile duct(11, 12).



**Figure 5:** In this figure, the papilla (arrow) can be clearly seen with no significant intra-duodenal component of the sphincter muscle. This flat papilla means only a very short sphincterotomy can be cut safely



**Figure 6:** By contrast, in this figure, a longer section of sphincter muscle is visible. The area between the arrows may be safely divided, leaving a much larger orifice for stone extraction

In cases where stones cannot be extracted with sphincterotomy alone, mechanical lithotripsy may be used to physically break up stones. This process involves capturing the stone within a basket, then using the basket itself to crush or fracture the stone within the bile duct to facilitate extraction(12). This is time-consuming, however, and involves increased risks of cholangitis and pancreatitis from small debris within the duct. In addition it has the unique risk that if a stone is lodged within the lithotripsy basket and cannot be crushed, the basket will be unable to be retrieved, necessitating urgent major open surgery to remove the impacted basket.

Lithotripsy can also be performed using laser or electrohydraulic methods. These require specific cameras (choledochoscopes), which are passed into the bile duct itself to perform lithotripsy under direct vision(12). These instruments are fragile and expensive. Lithotripsy by this method remains time-consuming and still has the risk of cholangitis and pancreatitis from passage of small stone debris. The technique is not widely available, although there are referral centres that do offer choledochoscopy.

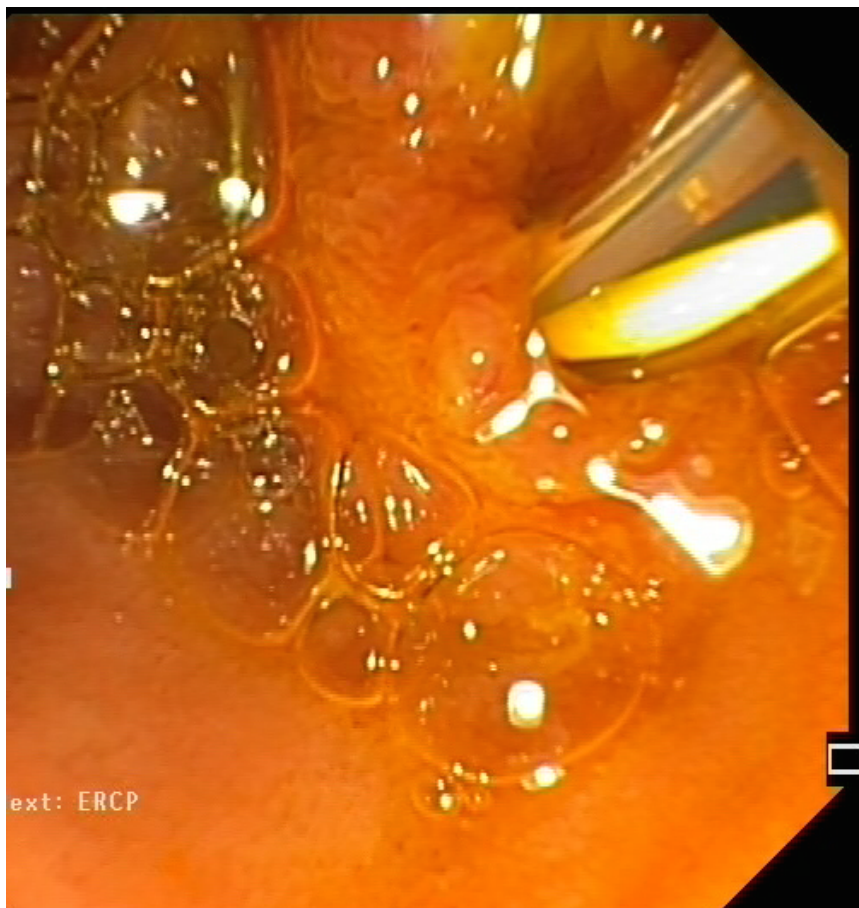
Finally, the traditional option for removal of large or multiple stones is that of surgical bile duct exploration. This remains an appropriate option in patients in whom endoscopic attempts have failed, or in whom the bile duct is inaccessible endoscopically due to previous surgery (e.g. gastrectomy). Surgery may be performed by either open or minimally invasive means and, prior to introduction of the technique of combined sphincterotomy + sphincteroplasty (see below), was the standard of care for extraction of large stones in our institution.

### **Sphincterotomy + large balloon sphincteroplasty – the combined technique**

To combat the above problems, the combined technique of sphincterotomy, followed by balloon sphincteroplasty was developed, first described by Ersoz *et al.* in 2003(8). Minor refinements have been described, but essentially the procedure consists of three steps:

1. Selective deep biliary cannulation of the bile duct in the standard fashion
2. Sphincterotomy, generally performed as a partial sphincterotomy i.e. less than the full extent of the biliary sphincter complex
3. Dilatation of the sphincter complex with a large dilating balloon, most commonly a controlled radial expansion-type balloon to a diameter, usually to 10mm, but may be more

These steps can be seen in the below images:

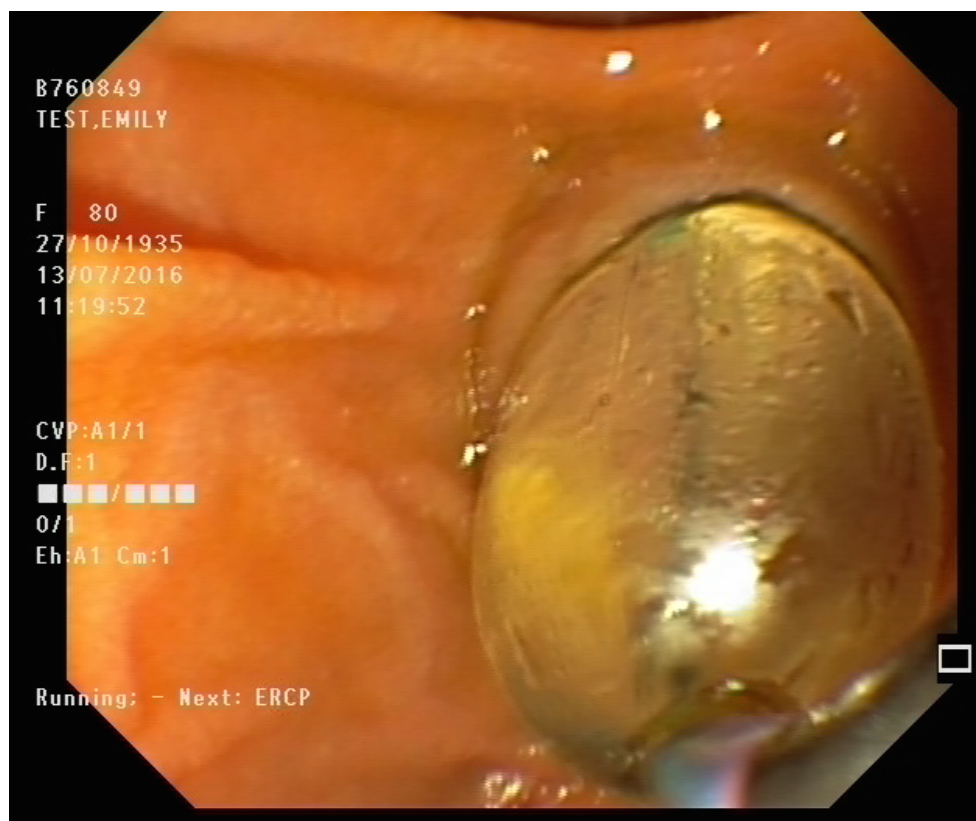


**Figure 7: Step 1: Selective deep biliary cannulation of the bile duct in the standard fashion**





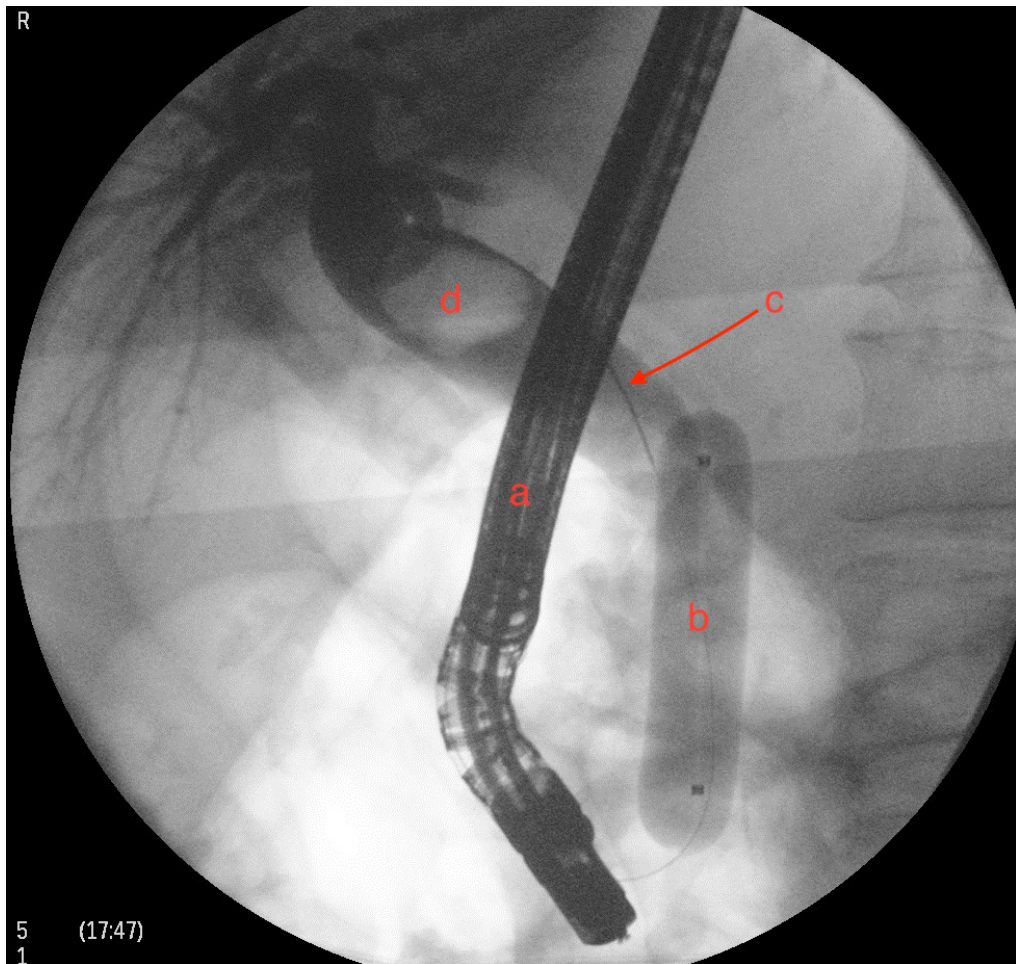
**Figure 8: Step 2: Partial sphincterotomy**



**Figure 9: Step 3: Balloon dilatation of the entire distal duct**

### **Mechanism of action**

The combined procedure is attractive in that the large dilatation allows enlargement of the sphincter to a greater extent than sphincterotomy alone. The balloon also dilates the whole distal common duct (see Figure 10). This means the whole distal duct and whole biliary sphincter complex is enlarged, rather than merely the very end of the sphincter muscle as is the case with sphincterotomy. In addition the combined technique may be technically simpler than cutting a very extensive sphincterotomy, especially in inexperienced hands.



**Figure 10: Image intensifier image of entire distal duct being stretched by dilating balloon.**

**a) duodenoscope, b) dilating balloon, c) wire in common bile duct, d) large stone in common hepatic duct**

Much of the concern with this combined technique relates to previous experience with sphincteroplasty as a primary technique, in particular the markedly increased risk of pancreatitis(10). However, these concerns do not seem to be borne out with studies of the combined technique, with similar rates of pancreatitis seen in the arms of trials comparing the combined technique to sphincterotomy alone(13). It is thought that the preceding sphincterotomy acts to direct the force of dilating balloon preferentially along the line of CBD and away from pancreatic duct, thereby decreasing the risk of pancreatitis(14).

### Comparing combined sphincterotomy + sphincteroplasty to sphincterotomy alone

Over 40 small studies demonstrate experience of institutions with the combined technique of sphincterotomy followed by dilatation. The key papers are summarised in **Figure 11**. Of these papers, three main meta-analyses attempt to summarise the available literature.

**Figure 11: Key sphincteroplasty papers**

Author	Type	Total pts	Combined technique success*	Complication rate	Complication measures used	Major issues with paper
Meine	Meta-analysis/ Review (retrospective)	1292	91%	5%	Total, individual complications e.g. bleeding perforation.  Definitions not specified	Largely small, uncontrolled case series.  Complication definitions not specified
Heo	RCT	200	83% vs 87% for sphincterotomy (ES) alone	5% vs 7% for ES alone	Total, individual complications.  Definitions based on Cotton criteria	Includes stones of all sizes -> includes patients not requiring the combined technique
Qian	RCT	132	81% vs 61% for ES alone $p = 0.046$	8% vs 12% for ES alone	Total and individual complications  Definitions similar to Cotton criteria	

Author	Type	Total pts	Combined technique success*	Complication rate	Complication measures used	Major issues with paper
Teoh	RCT	151	89% vs 89% for ES alone	7% vs 10% for ES alone	Total and individual complications,. Definitions similar to Cotton criteria	Includes small stones, patients with previous sphincter-otomy
Kim	RCT	55	85% vs 86% for ES alone	0 in both arms	Total and individual complications  Definitions similar to Cotton criteria	High use of mechanical lithotripsy in both arms (33%)
Feng	Meta-analysis (both RCTs and retrospective studies)	790	87% vs 84% for ES alone	6% vs 13% $p = 0.0007$	Total and individual complications definitions as per Cotton criteria	Multiple inappropriate trials included  Conflates prospective and retrospective data
Jin	Meta-analysis	621	82% vs 78% for ES alone	8% vs 11% for ES alone	Total and individual complications. Definitions not specified	

\*Success measured as extraction at first procedure

All  $p$  – values non-significant unless specified



Meine *et al.* have published the largest meta-analysis of available articles describing combined sphincterotomy and large balloon dilatation(15). This meta-analysis of 1292 patients in 21 articles suggests successful stone extraction of 91% after a single procedure, rising to 98% if multiple procedures were allowed. They describe a complication rate of around 5% with this technique. These figures compare very favourably with historical controls, e.g. initial extraction rates of 80.9% and complication rates of 10% in Weinberg's large meta-analysis of sphincterotomy alone (10)).

Meine *et al.* is the most comprehensive collection of the available literature at the time of writing (2014), covering almost all sizeable available series. Nevertheless, this data comes predominantly from uncontrolled retrospective case series, making comparing data between this combined technique and historical controls problematic. Most of the series also have fairly restrictive inclusion criteria and there will always be selection bias in a situation where the procedure is used on a discretionary basis – the endoscopist will be more likely to offer the novel technique in a patient perceived as more suitable. In particular, many studies specifically exclude patients with a higher risk of complications (e.g. multiple pancreatic cannulations, need for precut/needleknife access to the bile duct). As such, it is more difficult to generalise to real-world practice from this select subgroup.

The largest randomised controlled trial comparing sphincterotomy alone to a combined approach was that of Heo's group in Korea 2007(16). This randomised 200 patients with stones of any size to either sphincterotomy or combined sphincterotomy + balloon dilatation. It showed comparable efficacy in removal of stones, with no increase in complication rate. However, the study included all bile duct stones, rather than selecting the larger stones that may not be removable by sphincterotomy alone – for this reason, any benefit of the technique in dealing with larger stones (the topic under investigation in this thesis) will be diluted by the presence of small stones in each group. Importantly though, the identical complication rate in both arms confirms the safety of the approach.

Qian *et al.* randomised 132 patients, all of whom had stones  $\geq 15$ mm in size. 63 participants received a partial sphincterotomy, followed by a balloon sphincteroplasty of between 12 to 20mm as required(17). 69 participants received standard complete sphincterotomy. They found whilst overall stone clearance was ultimately the same between arms, clearance in the first session was significantly improved in the combined approach (80.9%) compared with standard sphincterotomy (60.8%) –  $p = 0.046$ . There was no significant difference in complications.

Teoh *et al.* describe a randomised controlled trial where 156 participants were randomised to either sphincterotomy alone or partial sphincterotomy + balloon

sphincteroplasty to a maximal 15mm(18). Whilst they found no difference in stone clearance between arms, mechanical lithotripsy was required in 46.2% of patients in the sphincterotomy alone group, significantly more than those using the combined technique (28.8% -  $p = 0.028$ ). Interestingly, Teoh *et al.* chose to include stones of all sizes, instead making a dilated bile duct ( $\geq 13\text{mm}$ ) the entry requirement. Nevertheless, they describe almost half of participants having a stone  $\geq 15\text{mm}$  (range 4 – 40mm), meaning it is likely they are predominantly dealing with large stone disease. In addition, a third (52/151) of patients included have had previous ERCP and/or sphincterotomy, a group who are specifically excluded from our patient population of interest.

Kim *et al.* randomised 55 participants with stones  $\geq 15\text{mm}$  to either sphincterotomy alone or partial sphincterotomy then dilatation to between 15 and 18mm(19). Unlike the other mentioned RCTS, they failed to show any difference between stone extraction or mechanical lithotripsy rates

There are two meta-analyses of available comparative studies on the topic, of varying quality. Feng *et al.* published a 2012 meta-analysis purporting to compare the available RCTs comparing sphincterotomy to techniques involving large balloon dilatation(20). The methodology of this meta-analysis is flawed unfortunately, as most of the studies included do not answer the question at hand. Firstly Heo's trial is included(16), which includes stones both large and small as described above. Secondly, it includes Stephanidis' trial of combined sphincterotomy + balloon dilatation vs sphincterotomy and mechanical lithotripsy(21), where all patients in the control group receive mechanical lithotripsy to crush their stones after sphincterotomy. Equating this control group to other "sphincterotomy alone" controls biases both the mechanical lithotripsy rate and complication rate in the sphincterotomy alone group. Thirdly, it includes both Itoi(22) and Kim's(23) non-randomised studies, which compares their results to historical controls. Finally, it includes Lin *et al.*'s study which compares sphincterotomy to balloon dilatation alone(24). In essence the meta-analysis attempts to group and compare interventions and studies that are not equivalent, making its conclusions meaningless.

Perhaps the most comprehensive review of the *relevant* comparative trials comes from Jin *et al.*'s 2014 meta-analysis(13). This compares five randomised controlled trials with 621 participants comparing sphincterotomy with techniques involving large balloon dilatation for large stones only ( $\geq 10\text{mm}$ ). Four of these five trials compared the combined technique to sphincterotomy alone. However, this meta-analysis includes Oh *et al.*'s trial of sphincterotomy vs balloon dilatation alone(25) and its relevance is diminished by this fact.

Jin *et al* describes equivalent efficacy in terms of ductal clearance both in the first session (82.2% vs 77.7% for sphincterotomy alone  $p = 0.17$ ) and overall (93.7% vs 92.5%,  $p = 0.54$ ). They did however suggest that use of a technique involving large balloon dilatation significantly decreased the need to resort to mechanical lithotripsy (15.5% vs 25.2%,  $p = 0.003$ ). In addition, there was no difference in complication rate between the two techniques (7.9% vs 10.7%,  $p = 0.25$ ). Thus, whilst its relevance is diminished by the inclusion of Oh *et al*'s irrelevant trial, it retains relevance by providing support for the use of a combined technique being at least as efficacious and at least as safe as sphincterotomy alone, for removing large stones.

### **Comparison to mechanical lithotripsy**

Mechanical lithotripsy involves mechanically crushing bile duct stones into fragments to facilitate their removal. This is a technique we resort to when the preferred technique of extraction of whole stones has failed. Stefanidis *et al*. performed a prospective, randomised controlled trial to compare sphincterotomy + balloon dilatation to sphincterotomy + mechanical lithotripsy in cases of large stones (12mm or greater)(21). This found no significant difference in efficacy of stone removal for combined sphincterotomy/sphincteroplasty vs sphincterotomy and mechanical lithotripsy. There was, however, a significantly higher rate of overall complications for mechanical lithotripsy (20% vs 4.4%,  $p = 0.049$ ), explained by a markedly increased rate of post-ERCP cholangitis (13% vs 0%,  $p = 0.026$ ). This increased complication rate confirms mechanical lithotripsy to be an inferior technique, meaning it is appropriate to consider using mechanical lithotripsy a marker of procedural failure.

## **Immediate vs delayed biliary dilatation**

Whilst the combination of sphincterotomy and sphincteroplasty is now well described, concern remains amongst many endoscopists about the appropriateness of combining the dilatation procedure together with sphincterotomy in the same session and that is the focus of this trial. There are two concerns here.

Firstly, there is a group of patients for whom minimising anaesthetic and procedural time is a high priority. These are systemically unwell patients, often elderly, in septic shock with cholangitis. In these patients prompt biliary drainage is the key, with definitive stone extraction a lesser concern, meaning stent placement to facilitate drainage may be all that is appropriate. In addition, injection of large amounts of contrast during a prolonged attempt at stone extraction may be harmful, spreading already-infected bile higher into the intrahepatic ducts and potentially causing a “septic shower”. In these patients, the safest option is to achieve rapid drainage by placing a stent and returning at a later date, when the sepsis has settled, for definitive sphincteroplasty and stone extraction. These patients are not the focus of this trial.

Secondly and more importantly for this trial, it is unclear whether the technique itself of performing balloon dilatation directly following sphincterotomy increases the risk of complications. This particularly applies if access has been difficult, with prolonged cannulation attempts, need for precut or needle knife access. In addition, more prolonged ampullary manipulation could possibly be more detrimental in certain cases thought to be more prone to complications, such as patients with active mild cholangitis or pancreatitis. When considering the results in the literature, it is important to note that many of the studies specifically exclude these patients, meaning their data comes from a carefully curated patient population, already expected to have good outcomes.

A method commonly performed in cases of failed stone extraction post sphincterotomy is placement of a biliary stent and return for large balloon sphincteroplasty at a later date. There are a number of potential benefits to this approach. Firstly, using a staged procedure prevents excessively prolonged manipulation at any single procedure (a known risk factor for post-ERCP pancreatitis)(26). Secondly, stenting and returning may allow a period of healing and remodelling of the sphincter complex, meaning that the action of the dilating balloon is performed on remodelled tissue, rather than a freshly cut sphincter. This theoretically may mean both a lower risk of perforation and bleeding from the large balloon dilatation and indeed a more effective dilatation with the force more effectively directed up the bile duct. Finally, stent placement itself may make subsequent retrieval attempts more efficacious by decreasing size and number of stones(27) (presumably by the friction of stent against stone eroding the latter).

Studies directly comparing the two approaches are limited to two retrospective series. The first is an English retrospective series presented in abstract only in 2015. It describes 202 patients over six years at a single institution, undergoing large balloon sphincteroplasty following sphincterotomy. Of these, 46% underwent this as a two-stage procedure, with the choice of approach at the discretion of the endoscopist. They found no significant difference in efficacy or complication in those undergoing either an immediate dilatation or at a second procedure(28).

Our unit has also previously published a retrospective series of 134 patients undergoing balloon sphincteroplasty, of whom 24% underwent dilatation immediately after their sphincterotomy(29). Again, the choice of immediate or delayed sphincteroplasty after sphincterotomy was at discretion of the surgeon involved. There was no statistically significant difference shown, either in efficacy or complications between either delayed or immediate dilatation, albeit with small numbers. Whilst not statistically significant, there was a non-significant trend towards both more bleeding ( $RR = 3.1, p = 0.14$ ) and indeed more overall complications in the immediate sphincteroplasty group ( $RR = 2.2, p = 0.25$ ). In addition, those in the immediate dilatation group generally received a less extensive dilatation than those receiving dilatation as part of a two-stage procedure (mean 10.34 vs 11.73 ( $p = 0.019$ )), meaning we may be under-calling any difference in complications.

Both these previous studies are retrospective series. The choice of whether to perform the balloon dilatation during the same procedure as the sphincterotomy or delay is at the discretion of the proceduralist, opening both series to selection bias. In addition, the arms are not directly comparable, as patients in the delayed sphincteroplasty groups have by definition already undergone at least one previous procedure, whose results are not included within the studies. For the patients in the immediate dilatation groups, the complications of cannulation and sphincterotomy are bundled together in the same procedure, whereas in the delayed groups, these complications occur in a previous procedure, which is not reported on. This biases both studies in favor of delayed dilatation. The more appropriate comparison would include all ERCP procedures in each patient's journey to duct clearance, not just those that include dilatation.

## **Efficacy measurement**

The most commonly described measure of efficacy in the ERCP stone extraction literature is extraction rate in first session. This is a less relevant endpoint for our study where one arm deliberately avoids stone extraction at the first session. In addition, the endpoint can be

manipulated by using secondary techniques (e.g. mechanical lithotripsy) to remove stones after failure of the trial technique.

Overall duct clearance rate is also a commonly used measure of efficacy, albeit as one would expect, most trials report very high overall extraction rates given enough attempts(15). This measure is highly unlikely to show a significant difference between arms in our study.

From a patient perspective, these larger stones have traditionally always required multiple procedures and the number of procedures required for clearance is an obvious marker of efficacy. This can be described both as the average number of procedures required and also the proportion of patients able to have their stones cleared by a single session – a clear advantage over our previous practice. Given the difficulties with other measures of clearance, this is the most clinically relevant measure of procedural efficacy in this case.

Mechanical lithotripsy rate is used in the ERCP literature as a surrogate marker for failed extraction. As previously described, the need to resort to mechanical lithotripsy suggests inadequate enlargement of the sphincter to allow the extraction of stones whole, the preferred technique. Use of mechanical lithotripsy is associated with a higher rate of post-ERCP complication, in particular cholangitis from incompletely drained small stone fragments(21). Therefore mechanical lithotripsy rates are used as a surrogate marker for failure of stone extraction.

It should be noted that while lithotripsy rates are commonly quoted in the literature as a surrogate marker of failure, our (otherwise busy) unit does not have the caseload to justify the purchase of the equipment to perform cholangioscopic-guided lithotripsy. Patients felt to require these procedures are referred to another centre for a subsequent procedure. In essence patients who might otherwise receive lithotripsy at other centres, at our centre are stented and return for a second procedure. This means that whilst lithotripsy rates are an appropriate surrogate for procedural failure, for practical purposes, this is an irrelevant metric for our particular institution.

Finally, procedural time is another surrogate measure for efficacy. Faster duct clearance leads to more efficient use of endoscopy time, allowing more procedures to be performed and potentially translating to cost savings. Itoi *et al.* report significantly shorter mean procedural time when a combined technique was used compared to sphincterotomy alone (32 vs. 40min,  $p < 0.05$ ), albeit in a non-controlled retrospective study(22).

## Complication measurement

Most trials report a raw figure for overall complication rate. This lumps together a disparate set of complications of varying severity and frequency. The major predictable complications of ERCP are bleeding, perforation, pancreatitis and infection.

### Cotton criteria

In 1991, a workshop of 25 experienced endoscopists met in an attempt to provide guidelines for prevention and management of complications. As part of this, a grading system was proposed, to allow standardised description and grading of severity of complications. Commonly described as the “Cotton Criteria” in honour of the primary author, this paper suggests a three-tiered grading of the complications bleeding, perforation, pancreatitis, cholangitis and basket impaction and is commonly used to provide standardised description of these complications in studies describing outcomes from ERCP(30).

**Figure 12: Cotton Criteria (adapted)**

Complication	Mild	Moderate	Severe
<b>Bleeding</b>	Clinical bleeding.  Haemoglobin drop <3g without transfusion	Transfusion ( $\leq 4$ units), no angiographic intervention or surgery	Transfusion $\geq 5$ units or intervention (angiographic or surgical)
<b>Perforation</b>	Very slight leak of contrast, treatable by fluids and suction $\leq 3$ days	Definite perforation treated medically between 4 and 10 days	Medical treatment for >10 days or intervention (percutaneous or surgical)
<b>Pancreatitis</b>	Clinical pancreatitis, amylase >3 times normal at >24 hours post-procedure, admission 2-3 days	Pancreatitis requiring admission 4-10 days	Hospitalisation for >10 days, or local complication or intervention
<b>Cholangitis</b>	>38°C  24-48 hours	Septic illness requiring >3 days of hospitalisation non-surgical intervention	Septic shock or surgical intervention required

NB: any intensive care admission grades as severe(30). Full criteria in appendix B.

There are some limitations to these “Cotton criteria”. In particular, the reliance on hospital length of admission for defining pancreatitis and for grading severity has been criticised as being able to be artificially manipulated, and may not reflect severity as well as other common measures, such as the Atlanta classification of pancreatitis or Clavien-Dindo score(31).

Nevertheless, these criteria are well accepted and used as standard nomenclature in the literature surrounding ERCP, just as they are for this study. This allows ready comparison of each of the individual complications with other studies available.

### **Comprehensive Complication Index**

The difficulty with comparing complications for ERCP (as it is with many surgical procedural trials) is that the complications are relatively uncommon and widely varying in severity. If comparing raw complication numbers, the majority are clinically mild pancreatitis, requiring admission and analgesia, but nil else. These are clearly less concerning than an increased rate of life threatening bleeding or perforation.

The Comprehensive Complication Index is a novel scale specifically designed to provide a grading system to measure severity of complications arising in surgical trials. It was first described in 2013 and calculates a weighted sum of all post-operative complications using a modification of the well-accepted Clavien-Dindo classification of complications (see appendix C). It then ascribes for each patient a score between 0 (best) – 100 (worst) allowing a graduated assessment of complication severity(32).

We postulate, similar to other surgical trials, that the Comprehensive Complication Index may be used as a novel technique to compare complications of varying severities in a more nuanced fashion. There are advantages in using this in comparing ERCP-related complications. Firstly, it provides a weighting to compare complications of different severities. Secondly, it makes allowance for multiple complications occurring to the same patient. It has been validated against both patient and physician rating of single and multiple complications, correlating closely with patient ratings of complication severity. In addition it has been shown to closely correlate with negative post-operative health status(32).

It must be stressed that this index has not been previously used in the context of endoscopic complications. It has, however been validated against results of multiple surgical RCTs, showing improved discriminatory ability compared to raw complication data alone(33). Following this, it has become accepted in the surgical complication literature,



being used as an endpoint to compare complications in over 60 surgical papers over the last four years.

## Methodology

### Hypothesis

That removal of large bile duct stones at ERCP with endoscopic sphincterotomy is performed more appropriately by immediate balloon sphincteroplasty rather than delaying balloon sphincteroplasty to a second procedure.

### Sample size calculation

Sample sizes were calculated using G\*power 3.1(34) based on calculations from our previously published series(29). We assumed an  $\alpha$ -error of 0.05, with a power ( $1-\beta$ ) of 0.8.

For the purposes of efficacy, we assumed a Chi-square test using conservative estimates for proportions of success in each arm (see Figure 13 below). These efficacy values are based on our previous clinical experience and clinically published series (see Appendix D), using more conservative estimates for success, particularly in the immediate dilatation arm. This estimates a required sample size of 14 participants to demonstrate superiority.

**Figure 13: Estimated proportional success for each arm**

No. procedures required for duct clearance	Immediate dilatation arm	Control
1	0.4	0
2	0.4	0.8
3	0.1	0.1
4	0.1	0.1

For the purposes of safety, our estimates were based on the complication rates seen in our previous series(29). These estimates will admittedly underestimate the complication rate in the control group, as each of these patients will have had a previous ERCP, the complications of which are not included in this data (see Appendix D for complication estimates). Bearing this in mind, we estimated a complication rate of 6% (6/104) for the control arm, compared with a complication rate of 13% (4/32) for an immediate dilatation approach. Using a Chi-square with one degree of freedom, we estimated 190 participants

would be required to show superiority assuming these rates, although it is almost certainly higher.

Not all complications are equal, particularly for ERCP, which has rare but potentially catastrophic complications such as duodenal perforation. Although clinically very relevant, an increase in rate of these rare but life-threatening perforations will not be detected unless a trial has a huge number of participants. Perforation rate in large real-world series such as the Gallriks compulsory database of all Swedish ERCPs suggest a perforation rate of 0.3%(6). If the perforation rate was doubled by immediate dilatation technique under trial, over 2,600 patients would be needed to show statistical significance. Even if the perforation rate were increased five-fold by using the immediate dilatation approach, 164 patients would still be required to show statistical significance. Hence it is not possible, in practical terms, to perform a trial that could capture increased rates of this rare but serious complication.

Our previous series had suggested a possible difference in a single individual complication – bleeding. Immediate dilatation was suggested to be associated with a three-fold increased risk of bleeding(29), although the small numbers concerned meant this did not reach statistical significance. This estimate, on a per-procedure analysis, suffers from the same issue of that for overall complications – it misses the complications from the preceding ERCPs for patients in the control arm. Nonetheless, using data from this series, we estimated a bleeding rate of 3% in the control arm compared with 9% in the immediate dilatation arm. Assuming this three-fold increase to be correct, we performed a sample size calculation, estimating 180 participants would still be required to confirm superiority, even though the true number required is likely significantly higher.

Remembering that our trial can only include a relatively small subset of ERCP patients with large stones, we can see that for both raw complication number and individual complications, estimated sample sizes are prohibitively large. Even for a high-volume unit like ours, it would take at least 5 years to accrue the necessary numbers for comparison of complications, bleeding or perforation. Hence our analyses of both overall and individual complication rates are, in effect, predominantly exploratory outcomes only.

## **RCT design**

A non-blinded randomised-controlled trial was performed. Participants were recruited from two high-volume tertiary referral ERCP centres.: Box Hill Hospital and St Vincent's Hospitals, both in Melbourne, Australia.

### **Inclusion criteria**

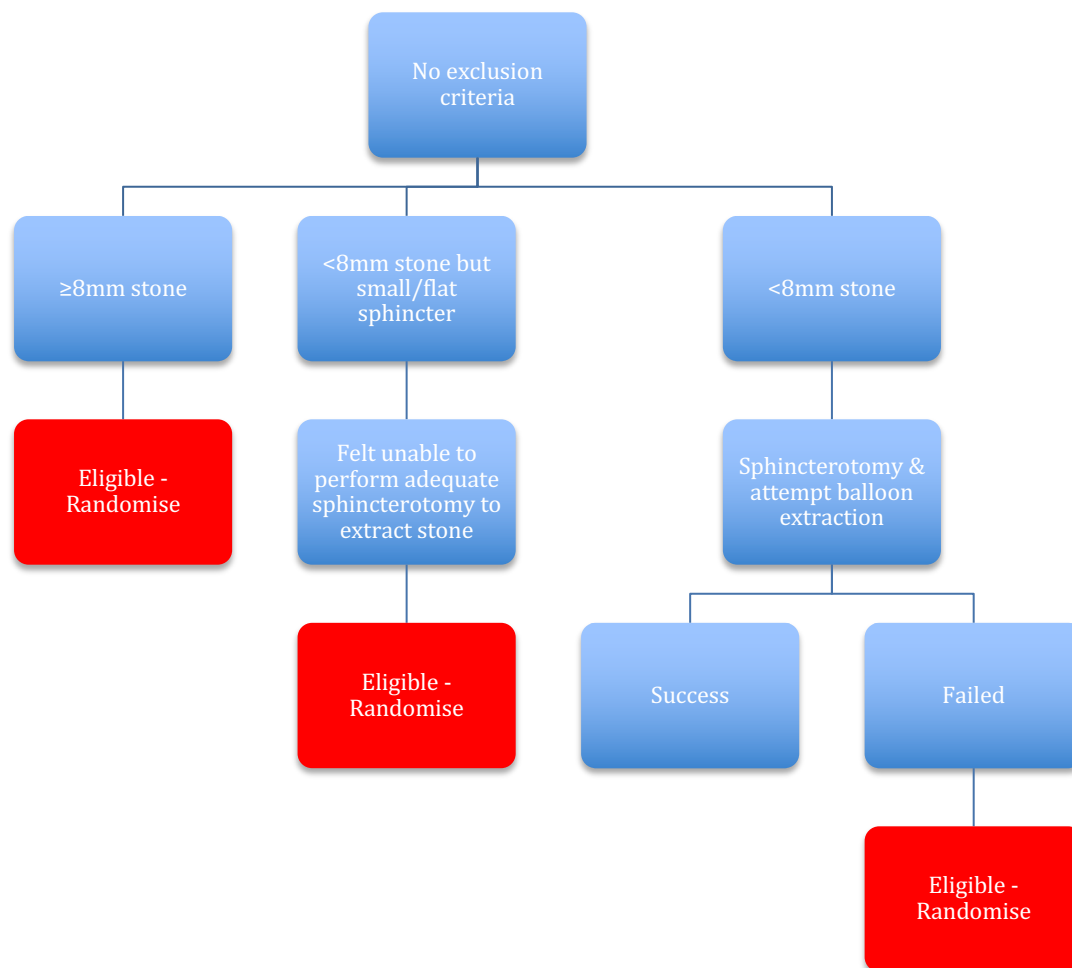
Patients were screened if they were undergoing ERCP for potentially large bile duct stones, and had never had sphincterotomy before. This was determined from available pre-procedural imaging (generally ultrasound or MRCP) suggesting common bile duct stones. Given the unreliability of preoperative imaging (particularly ultrasound) at determining stone size, patients who were preoperatively thought to have smaller stones were also consented to the *possibility* of being in the trial, should their stones be unexpectedly unable to be extracted by conventional means.

A subset of these patients were found to be eligible for the trial, based on intraoperative findings measured against the 12mm duodenoscope size at cholangiogram. These patients had stones deemed too large to extract with sphincterotomy alone. This was determined if either:

- a) Stone  $\geq 8\text{mm}$
- b) Stone  $< 8\text{mm}$  but the sphincter was too small to allow a sufficiently large sphincterotomy to extract the stone
- c) Stone  $< 8\text{mm}$  but had failed stone extraction after maximal sphincterotomy

These intra-operative criteria for trial eligibility mean that significantly more patients were approached and consented prior to their procedure than were eventually randomised, to ensure the maximum potential participants were captured.

**Figure 14: Inclusion criteria diagram**



### Procedural Details

All participants had an ERCP performed using an Olympus 12mm duodenoscope. Deep biliary cannulation was obtained using either a Jagtome (Boston Scientific) or Needle-knife papillotome (Boston Scientific). A cholangiogram was performed and the stone size measured against the size of the duodenoscope (12mm) using image intensifier. If participants were deemed to be eligible, a partial sphincterotomy was cut according to the endoscopist's assessment of the papilla and randomisation was performed. Needle papillotome or precut sphincterotomy were allowable as long as selective biliary cannulation was obtained.

ERCP was performed by either an experienced hepatobiliary surgeon with extensive ERCP experience, or by a senior HPB surgical fellow training in ERCP under direct consultant supervision. There was direct trainee involvement in most cases. Five hepatobiliary consultants either performed or directly supervised the procedures.

Participants who were unable to speak English were consented using the hospital's in-house interpreters, or a professional phone interpreter service. If a patient was unable to consent for him or herself, a proxy provided informed consent both for the procedure (i.e. just as in the non-trial situation) and for the trial.

### **Exclusion criteria**

Predominantly patients who were excluded were those who were felt to be too sick for a prolonged procedure. That is, patients with septic shock, severe pancreatitis or cholangitis. In these patients it is generally possible to place a stent without the need for stone extraction, ensuring a short anaesthetic with effective establishment of biliary drainage – this represents the standard care for patients in this scenario. This was determined by the clinical judgement of the treating surgeon in each case. In addition patients with unacceptably high risk of bleeding were excluded – those with a coagulopathy (International Normalised Ratio (INR) > 1.5, platelet count <50,000/ $\mu$ L) or on anticoagulation or strong antiplatelet drugs – aspirin was admissible as previous reviews have shown this not to be associated with increased bleeding risk(35).

Patients with severe co-morbidities, who were felt to be at increased risk from multiple anaesthetics, were also excluded. This decision was up to the clinical judgement of the treating surgeon and anaesthetist in each case. Patients with severe co-morbidities were treated with a combined procedure (sphincterotomy and then immediate sphincteroplasty) on the basis that this approach has been part of the treating unit's therapeutic options(29), and that it was anticipated to deal with the stones at a single procedure. This approach did mean that the participants most likely to be harmed by a delayed approach were excluded from the trial. This does potentially bias the results against the intervention treatment (immediate dilatation) arm, but it was felt to be ethically inappropriate to enrol such patients.

### **Randomisation procedure**

Randomisation was performed using a closed envelope technique, using a computer-generated random series, 1:1, block-randomised in groups of 20. Randomisation took place once the cholangiogram had been performed, if the patient met inclusion criteria.

## **Arms**

Upon performing the cholangiogram, the presence of stones was confirmed, and stone size measured. Having confirmed their eligibility, participants would then be randomised to one of two arms:

### **Intervention – Immediate dilatation arm**

Following sphincterotomy, large balloon sphincteroplasty was performed with a dilating balloon (either 10mm Hurricane RX biliary balloon – Boston Scientific, or wire-guided controlled radial expansion (CRE) 12 or 15mm balloon – Boston Scientific). Balloon position was confirmed with image intensifier.

Dilatation occurred in a stepwise fashion, each step for 60 seconds – 10mm, then 12mm, then 13.5, then up to 15mm depending on the necessary size of dilatation. The dilatation was taken to whichever was the smallest of; 15mm; the maximal size of the bile duct; or 2mm greater than the diameter of the largest stone.

An attempt to extract the stones was then performed using an extraction balloon (9-12mm or 12-15mm, Extractor Pro, Boston Scientific). If extraction was unsuccessful, a 5cm 7F double pigtail plastic biliary stent was placed.

### **Control – stent and delayed sphincteroplasty arm**

If randomised to control, following sphincterotomy, a 5cm 7F plastic double pigtail biliary stent was placed. No attempt was made to extract the stone at this stage.

The participant would then return for a repeat ERCP 4-6 weeks later, at which stage the stent was removed and balloon sphincteroplasty performed using the same protocol as above. If stone extraction at this stage was unsuccessful, a 5cm 7F double pigtail plastic biliary stent was again placed.

Essentially the difference in the two arms is timing – in the delayed sphincteroplasty group, 4-6 weeks occurs between sphincterotomy and balloon dilatation, allowing time for sphincter healing and remodelling. The components of the treatment otherwise remained the same.

## **Antibiotic and NSAID protocol**

Antibiotics and rectal non-steroidal anti-inflammatories (NSAIDs) are used during ERCP to decrease the risk of cholangitis and pancreatitis respectively. Patients received antibiotics and rectal NSAIDs selectively, as per our institutional protocol (see appendix A).

Specifically, patients undergoing their first ERCP did not receive antibiotics unless they had active infection, or the bile duct was unable to be drained. Piperacillin-Tazobactam 4.5g was our first-line antibiotic, administered on induction. Patients undergoing repeat ERCP received a single dose of prophylactic antibiotic due to the risk of instrumenting an already colonised bile duct.

As per our institutional protocol, we use a selective approach to prophylactic NSAIDs for prophylaxis against post-ERCP pancreatitis(36). Patients undergoing their primary procedure received 100mg rectal indomethacin at the completion of the procedure, as did patients undergoing a subsequent procedure if they had risk factors for post-ERCP pancreatitis. These risk factors included difficult or prolonged cannulation, injection of contrast into the pancreatic duct (pancreaticogram), multiple pancreatic wire-cannulations, young female patients and previous post-ERCP pancreatitis.

Indomethacin was omitted in cases of allergy, significant renal impairment, NSAID-induced asthma or significant bleeding risk.

Pancreatic duct stents were not placed prophylactically to prevent post-ERCP pancreatitis, instead relying on pharmacological prophylaxis with NSAIDs as described above.

## **Measurement of Complications:**

Peri-procedural complications were documented during the first week post operatively, either during their inpatient admission, or if at home, via a follow-up phone call. All patients at our centre receive a follow-up phone call at one week, to accurately audit outcomes and allow comprehensive identification of perioperative complications, particularly in patients sent from rural centres, who may be otherwise missed.

Standard post-ERCP complications were measure using the standard grading scale: Cotton's criteria(30), namely pancreatitis, cholangitis (infection), perforation and bleeding. This measures the above complications each on a standardised 3-point scale (see appendix B).



### **Comprehensive Complication Index**

The difficulty with comparing complications for ERCP (as it is with many surgical procedural trials) is that the complications are relatively uncommon and widely varying in severity. Moreover, mild problems make up the majority of complications, whereas serious and/or life-threatening complications are rare. In this way, the lesser complications tend to dilute out the serious ones (which are of greater interest). The Comprehensive Complication Index is a novel scale specifically designed to provide a grading system to measure severity and number of complications arising in surgical patients(32).

We measured severity using the Comprehensive Complication Index across the gamut of complications to provide a means of grading and comparing complications of differing severities. Complications were documented prospectively and participant medical records re-examined retrospectively to collate all complications and their management throughout their admission. All complications were entered anonymously into the online CCI calculator to calculate Comprehensive Complication Index scores for each patient(37). This was used as a novel technique to try to compare complication rates between the two arms and provide a more logical method to weigh disparate complications such as perforation and pancreatitis against each other.

It is important to note that this has not been validated in the context of endoscopic intervention, and the use of this index in this trial does represent the first report of the Comprehensive Complication Index in an endoscopic procedure. This was used in this study in an exploratory fashion, but may become a useful tool for comparing disparate and rare complications of endoscopy, just as it has in the wider surgical complication literature.

### **Data collection**

Data was collected peri-procedurally, during inpatient stay (for elective cases usually a day procedure) and at routine follow-up. This routine follow-up consisted of either a phone-call by primary investigator or review in outpatient clinic or both, depending on clinical appropriateness. In most cases phone call follow-up was performed within 1-2 weeks, with clinic appointments usually made at four weeks.

These multiple collection points allowed accurate collation of data from both the procedure itself and the time-period within which procedural complications occur. Our routine post-operative follow-up ensured all complications occurring post-discharge were captured.

Prospective data collection was double-checked by the primary investigator against the hospital record to ensure other complications were not missed. All complication data was collected on a template, based on the Cotton criteria as a guide to definition and severity.

Procedural and anaesthetic times were documented by nursing staff as part of their routine documentation, independent to the trial. Stone and bile duct size were calculated based on the appearance of the cholangiogram at time of ERCP, comparing stone and bile duct measurements to the baseline size of the duodenoscope (12mm).

### **Analysis techniques**

Statistical analysis was performed using SPSS version 24.0.

Categorical data were compared for significance using Fisher's exact test for binary outcomes and Chi-square analysis for non-binary or ordinal outcomes.

Parametric numerical data were compared for significance using one-way Anova or Student T-test.

Non-parametric numerical data were compared using Mann-Whitney U-test, with medians compared using Mood's median test in cases of data with extreme outliers (such as radiation dose).

Univariate analysis for correlation was performed using Pearson's coefficient.

A *p*-value of <0.05 was considered statistically significant.

Sample size calculations were performed using G\*power 3.1(34) using the techniques described in relevant section above.

## Results

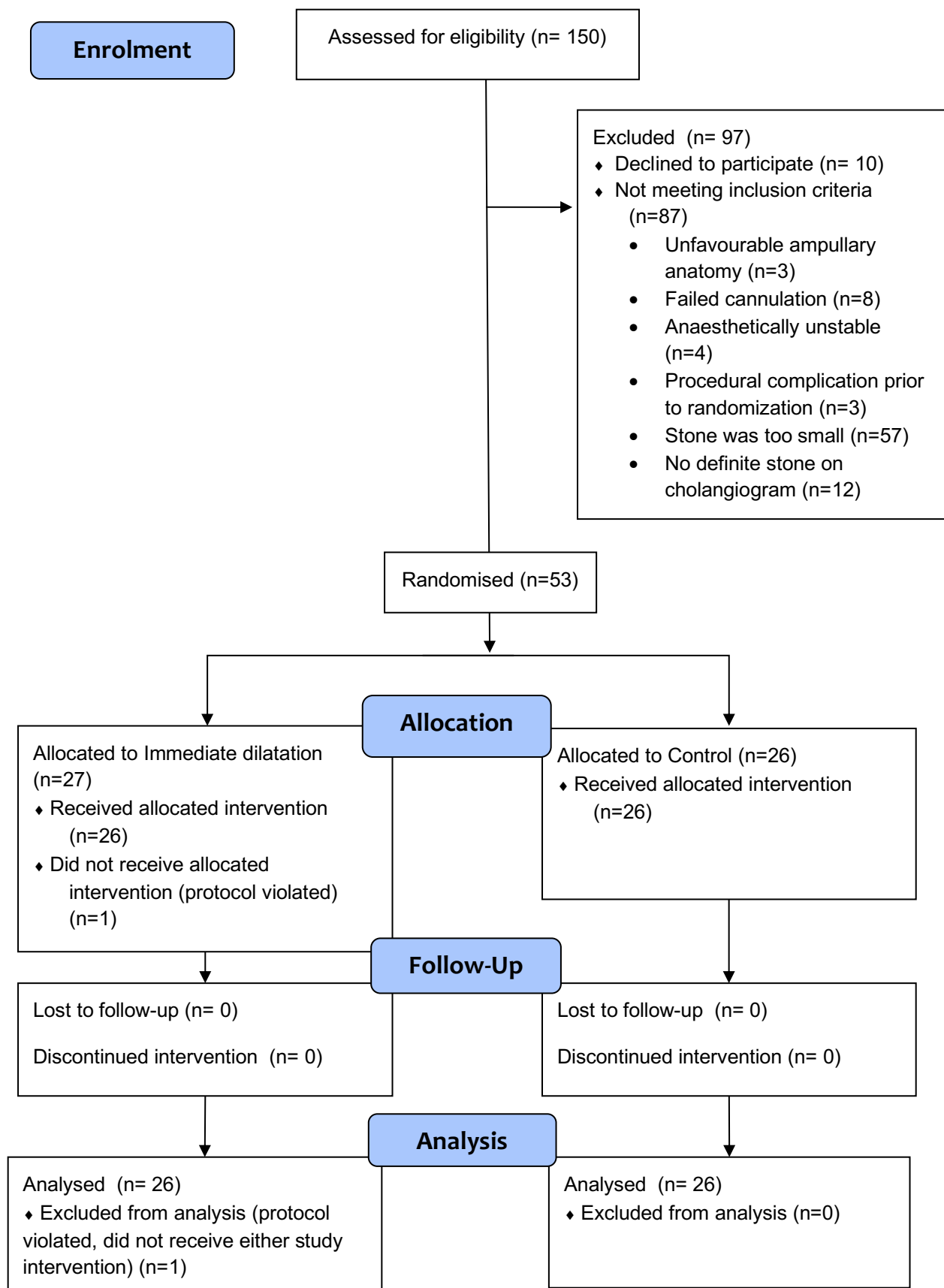
### Enrolment

In total 150 patients undergoing their first ERCP for potentially large stones were approached. In line with the usual acute presentation of choledocholithiasis, only three of these were electively booked outpatients with the rest presenting on an urgent or semi-urgent basis. Of these 150 potential participants approached, 10 declined to consent to the trial. Of the 140 consented patients, 53 were found to be eligible at time of randomisation (i.e. after cannulation and performance of cholangiogram).

The remaining 87 underwent their ERCP but did not meet eligibility criteria at the time of cholangiogram. Three had unfavourable ampullary anatomy precluding sphincteroplasty. In eight patients cannulation was unsuccessful and the ERCP could not proceed. Four were excluded as they became anaesthetically unstable, precluding a prolonged procedure. Three had a procedural complication prior to randomisation. Of those who proceeded to cholangiogram, in 57 we found that the stone was too small for eligibility criteria and in 12 did not find a definite stone on cholangiogram.

Of the 53 eligible participants, in all but one participant the appropriate intervention was performed after randomisation. The sole exclusion here was a participant randomised to immediate dilatation, in whom the proceduralist broke protocol, performing neither of the trial procedures. This patient was censored from follow-up and analysis, leaving 52 participants for analysis. For completeness we also reanalysed all outcomes on an intention-to-treat basis including this censored patient and this did not significantly alter results for any outcome. All the remaining participants had the allocated procedure performed and none were lost to follow-up. This is summarised in the below CONSORT diagram.

**Figure 15: CONSORT diagram**



## Demographics

Demographic data was collected with regards to age, gender and comorbidities and were similar across both groups. These data are summarised in the following table.

**Figure 16: Demographic Data**

	<b>Immediate dilatation</b>	<b>Control</b>	<b><i>p</i>-value</b>
<b>Age (mean years)</b>	62.23	70.5	0.09 <sup>a</sup>
<b>Gender</b>	50%	50%	1.00 <sup>b</sup>
<b>ASA* (median)</b>	2	2	0.33 <sup>c</sup>
<b>Smoking currently</b>	32%	12.50%	0.17 <sup>b</sup>
<b>Cardiac disease</b>	20%	38%	0.22 <sup>b</sup>
<b>Diabetes</b>	16%	17%	1.00 <sup>b</sup>
<b>Aspirin</b>	23%	36%	0.37 <sup>b</sup>

**a = Fisher's exact test**

**b = Student T-test**

**c = Chi-square test**

**\*American Society of Anaesthesiologists score**

### Age

Mean age was 66.3 years, with a range of 21 – 95 years. This was similar in both groups – mean for the intervention group 62.2 years and 70.5 for the control group ( $p = 0.08$  using Student T-test).

### Gender

There was no gender predominance in either group – 50% of participants were male in each group ( $p = 1$  using Fisher's exact test).

### Smoker

There was no significant difference in smoking rate between each arm (Figure 17)

**Figure 17: Smoking status**

		Smoking Status			Total
		Current	Never	Ex	
Arm	Immediate Dilatation	8	12	5	25
	Control	3	15	6	24
Total		11	27	11	49

Comparing current smokers only,  $p = 0.17$  using Fisher's Exact Test

### Cardiac disease

29% of participants had known cardiac disease at time of their initial ERCP. There was no significant difference between arms (20% vs 38%,  $p = 0.22$  using Student T-test).

### Diabetes

16% of participants in the immediate dilatation group were diabetic, compared to 16.7% of the control group ( $p = 1$  using Fisher's Exact Test)

### Aspirin

Aspirin was the only blood thinning medication allowed as part of the study protocol, in line with our current practice for performing sphincterotomy or sphincteroplasty. Overall, 29.4% of participants were taking aspirin. There was no significant difference between arms in aspirin use (23.1% in the immediate dilatation arm vs 36% in the control arm) –  $p = 0.37$  using Fisher's exact test.

## ASA

ASA (American Society of Anaesthesiologists) score is a standardised measure of anaesthetic risk used almost universally by anaesthetists in their perioperative assessment. It encompasses both acute illness and comorbidities. It is used predominantly as a proxy for comorbidities.

**Figure 18: ASA (American Society of Anaesthesiologists) score**

- **Class 1** = A normal healthy patient
- **Class 2** = Mild systemic disease
- **Class 3** = A patient with severe systemic disease that is not incapacitating
- **Class 4** = A patient with an incapacitating systemic disease that is a constant threat to life
- **Class 5** = A moribund patient not expected to survive 24 hours with or without an operation(38)

All but three participants were ASA 2 or 3 – an expected result as the patient population with symptomatic choledocholithiasis generally is elderly, with some systemic disease. ASA 4 patients were generally excluded due to their instability. There was no significant difference between arms –  $p = 0.33$  using Chi-square test.

**Figure 19: ASA distribution**

		ASA				Total
		1	2	3	4	
Arm	Immediate Dilatation	2	13	9	1	25
	Control	0	13	12	0	25
Total		2	26	21	1	50

### Indication for initial ERCP

As per the inclusion criteria, all participants in this study had symptomatic bile duct stones as the indication for their initial ERCP. However there is a spectrum of diseases caused by bile duct stones. It was postulated that some of these might make ERCP and stone extraction more difficult or indeed predispose to complications.

These preoperative considerations are summarised in Figure 20:

**Figure 20: Indication for ERCP**

	Immediate dilatation	Control	P-value
Pancreatitis	15.4%	11.5%	1.00 <sup>b</sup>
Infection	26.9%	42.3%	0.38 <sup>b</sup>
Bilirubin (median, range)	33 (4-322)	27 (4-141)	0.58 <sup>d</sup>
Highest Bilirubin (median, range)	33 (4-322)	27 (4-191)	0.67 <sup>d</sup>

**b = Student T-test**

**d = Mann-Whitney U-test**

### ***Pancreatitis***

A common indication for ERCP is gallstone pancreatitis. Pre-existing pancreatitis may cause ampullary and duodenal swelling and distortion, leading to increased procedural difficulty. Furthermore, it may skew results for the most common post-ERCP complication: pancreatitis. With regards to pre-procedural pancreatitis, 13.5% had gallstone pancreatitis prior to ERCP. There was no statistically significant difference between the two arms (20% had pre-procedural pancreatitis in the immediate dilatation arm, compared with 8% in the control arm –  $p = 0.38$  using Fisher's exact test).



### ***Cholangitis***

Preoperative infection (cholangitis) may predispose to post-operative sepsis. In addition, it may make the ampulla swollen and friable, predisposing to bleeding. It is worth noting that patients with significant haemodynamic instability due to sepsis were excluded from this trial. Cholangitis was present in 34.6% of patients overall. This was not significantly different between arms (26.9% in the immediate dilatation vs 42.3% in the control arm) –  $p = 0.67$  using Student T-test.

### **Serum bilirubin**

We measured preoperative bilirubin as a measure of liver dysfunction and obstruction preoperatively. Biliary obstruction is associated with fat-soluble vitamin malabsorption and vitamin K-dependent factor coagulopathy, leading to increased risk of bleeding.

Two measures were taken. Firstly, we noted the most recent preoperative serum bilirubin, within 48 hours of ERCP. Median bilirubin was  $30\mu\text{mol/L}$  overall, with no significant difference between the immediate dilatation ( $33\mu\text{mol/L}$ ) vs the control ( $27\mu\text{mol/L}$ ) ( $p = 0.58$  using Mann-Whitney U test).

We also measured the peak bilirubin, being the highest serum bilirubin level during the acute illness leading to admission. Similar results were seen here: overall median bilirubin  $35\mu\text{mol/L}$ , with no significant difference between the immediate dilatation ( $33\mu\text{mol/L}$ ) vs the control ( $27\mu\text{mol/L}$ ) ( $p = 0.58$  using Mann-Whitney U test). For both measures, serum bilirubin ranged widely in both arms (4 -  $322\mu\text{mol/L}$ ).

On univariate analysis, both measures of bilirubin correlated with an increase in our primary endpoint – number of ERCPs required. The immediate preoperative bilirubin correlated more strongly (Pearson correlation 0.34,  $p = 0.016$ ) than the maximum bilirubin measurement (Pearson correlation 0.28,  $p = 0.048$ ).

### **Inclusion criteria**

As previously discussed there were three separate paths to inclusion. This was either:

- a. Stone  $\geq 8\text{mm}$
- b. Stone  $< 8\text{mm}$  but sphincter too small to allow sufficiently large sphincterotomy to extract stone
- c. Stone  $< 8\text{mm}$  but failed stone extraction after sphincterotomy

Despite multiple possible paths for enrolment, in essence almost all participants had stones  $\geq 8\text{mm}$ . A single patient was included after sphincterotomy and failed extraction of a 7mm stone. This patient had a flat papilla, with minimal intra-duodenal component, and this precluded a wide sphincterotomy. In addition, five patients fulfilled two inclusion criteria, having a flat papilla that would have precluded extraction, even had the stones concerned been  $< 8\text{mm}$  in size.

## **Procedural details**

### **Cannulation rates**

Of the 140 participants who were consented, successful cannulation occurred in 132 (94%). This compares favourably to international benchmarks e.g. Kapral *et al*'s study attempting to benchmark ERCP performance in Austria, demonstrating a 86% cannulation rate for ERCPs performed by high-volume ( $> 50$  ERCPs/year) endoscopists(39). This confirms appropriate technical performance of ERCP overall within our study.

### **Cholangiogram**

#### **Stone size**

The mean size was slightly higher in the control group – 11.2mm compared with 10.5mm in the immediate dilatation group which was neither statistically nor clinically significant ( $p = 0.36$  using Student T-test).

On univariate analysis (using Pearson's correlation) a correlation is seen between a stone size and number of ERCPs procedures required (Pearson correlation 0.322,  $p = 0.02$ . This correlates with other trials which suggest stone size is an independent predictor for difficulty of extracting stones(40).

#### **Duct size**

Similar to stone size, there was no significant difference in size of bile duct between the two groups. Mean maximal bile duct diameter was 12.4mm in the intervention compared with 12.9mm in the control group ( $p = 0.47$  using Student T-test). A participant's bile duct size did not correlate with the number of ERCPs they required.

**Antibiotic use**

Antibiotics were used on a selective basis to prevent post-ERCP cholangitis. As per our institutional protocol, antibiotics were used if the bile duct had been previously colonised (previous biliary intervention) or if there was possible biliary sepsis (see appendix A). This was a surrogate marker for suspected biliary sepsis, a predisposing factor for cholangitis following ERCP. There was no difference in antibiotic use at initial ERCP (46.2% in the intervention group vs 53.8% in the control) ( $p = 0.78$  using Fisher's exact test).

**Non-Steroidal Anti-Inflammatory (NSAID) use**

NSAIDs were used on a selective basis to prevent post-ERCP pancreatitis. As per our institutional protocol (see appendix A), rectal indomethacin was administered in the index procedure in all cases unless contraindicated. All but one patient undergoing sphincterotomy and immediate dilatation had rectal indomethacin administered in this first ERCP. Four participants in the control group did not receive rectal indomethacin. This difference was not statistically significant ( $p = 0.35$  using Fisher's exact test).

Indomethacin was not administered in four cases due to significant renal impairment and in the final case due to documented NSAID allergy.

### Sphincterotomy approaches

In cases where we are unable to access the bile duct directly with a wire, use of different techniques to cannulate are required. We measured the rate of resorting to precut sphincterotomy (freehand using the standard papillotome without wire cannulation) or needle-knife sphincterotomy (see

Figure 21). Both of these are associated with difficulty of cannulation and increased complications, namely pancreatitis, bleeding and perforation.

Needle-knife sphincterotomy was used in 11.5% of patients, identical in both arms. The technique of precut freehand sphincterotome access was also used in 11.5% of participants in both arms ( $p = 1$  using Fisher's exact test). The need to use precut techniques is somewhat higher in our series than that seen in the literature (12.4% of sphincterotomies in the Swedish nationwide ERCP audit, GallRiks(6)). This may reflect our practice of early use of precut techniques in difficult ERCPs, a practice shown to improve successful cannulation, with decreased risk of complications in experienced hands(41).

**Figure 21: Sphincterotomy Technique**

		Standard papillotome	Needleknife	Precut sphincterotome freehand	Total
Arm	Immediate Dilatation	20	3	3	26
	Control	20	3	3	26
Total		40	6	6	52

### Pancreatic cannulation rate

Inadvertent cannulation of the pancreatic duct during ERCP may be associated with an increased risk of post-ERCP pancreatitis. There was no significant difference in mean number of pancreatic cannulations in either the immediate dilatation (1.08) or the control (0.81) arm ( $p = 0.58$  using Student T-test).

In addition, in one participant in the immediate dilatation arm, an inadvertent pancreaticogram was performed. Contrast within the pancreatic duct system significantly increases the risk of pancreatitis(3).

**Figure 22: Procedural data**

	Immediate Dilatation	Control	<i>p</i> -value
Mean stone size (mm)	10.5	11.2	0.36 <sup>b</sup>
Mean CBD size (mm)	12.3	12.8	0.47 <sup>b</sup>
Antibiotic use	46.2%	53.8%	0.78 <sup>a</sup>
Rectal indomethacin	96.2%	84.6%	0.35 <sup>a</sup>
Needleknife sphincterotomy	11.5%	11.5%	1 <sup>a</sup>
Precut sphincterotomy	11.5%	11.5%	1 <sup>a</sup>
Pancreatic cannulations (mean)	1.08	0.81	0.58 <sup>b</sup>

**a = Fisher's exact test**

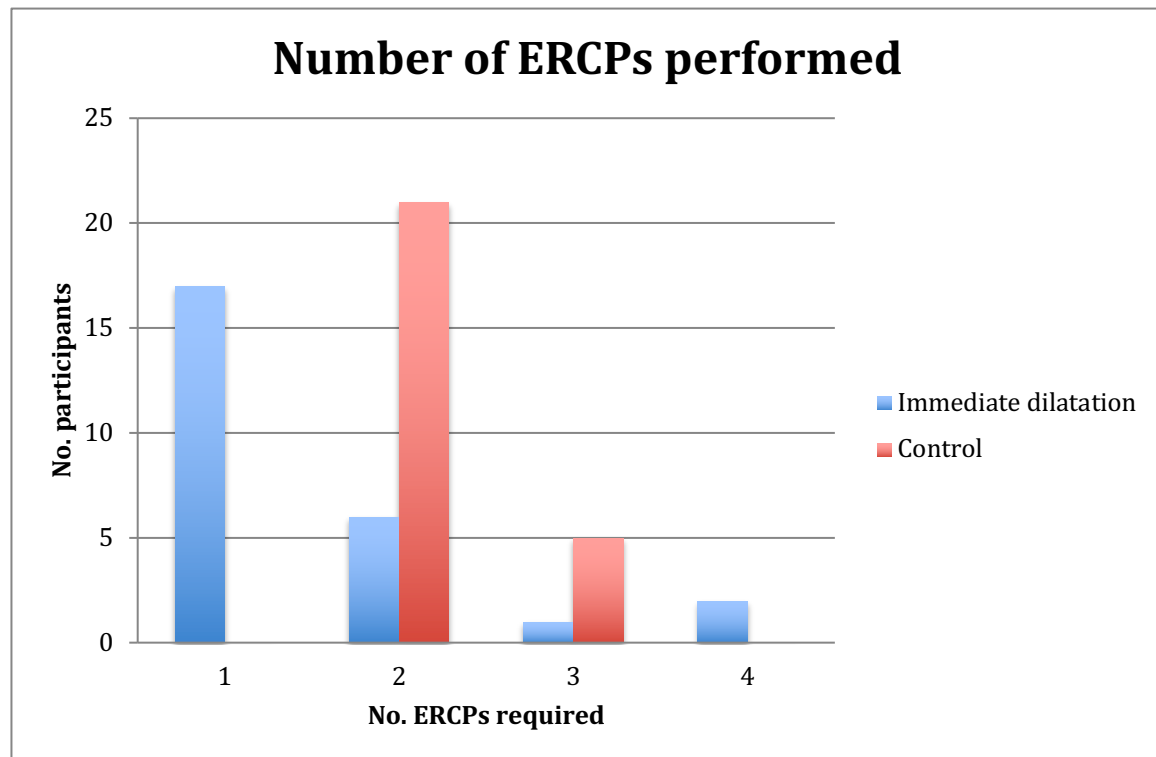
**b = Student T-test**

## Efficacy

### Procedures to clearance

The primary endpoint measured was number of procedures to clearance of all bile duct stones.

Figure 23: Number of ERCPs performed



Using the control group to represent our previous practice, we showed a significant decrease in the number of procedures required in the immediate dilatation group when compared to the control group –  $p < 0.001$  using Chi-square analysis. This confirms the statistical significance of this clinically significant improvement in care by an immediate dilatation approach.

For participants in the immediate dilatation group, 17/26 participants (65.4%) required only one procedure overall. Given that in our previous practice all of these patients would have required at least two endoscopic procedures, and in some cases surgical bile duct exploration this is a clinically significant improvement in practice. Clearance rate after a single ERCP can be compared between arms using Fisher's exact test and again this confirms statistical as well as clinical superiority ( $p < 0.001$ )

### **% Clearance**

88% of participants in the immediate dilatation group had a clear cholangiogram at the completion of their procedure. Of these, four patients had a stent left to facilitate drainage prior to cholecystectomy despite duct clearance. One required further ERCP to deal with a bile leak at cholecystectomy. Three participants had further duct stones found at cholecystectomy. Two of these participants had delayed their planned cholecystectomies for personal reasons.

Of the participants in the control group, 80% had a clear cholangiogram at the completion of their second procedure. Given these patients are just stented at their first procedure, without attempting stone extraction, 80% had their duct cleared at the first attempt. There was no significant difference for this metric between arms ( $p = 0.70$  using Fisher's exact test).

Overall clearance rate was higher in the immediate dilatation group, albeit not to a statistically significant extent. All patients in the immediate dilatation group had endoscopic duct clearance, albeit one needed referral for endoscopic cholangioscopy and electrohydraulic lithotripsy (3.8%). By contrast, three participants (11%) in the control group required surgical bile duct exploration for duct clearance ( $p = 0.24$  using Fisher's exact test).

### **Time**

Time as a measure of efficiency was measured in multiple ways. Firstly overall procedural time and in addition anaesthetic time overall were measured as a measure of efficiency of use of endoscopy resources. Admission duration overall is a function of multiple factors but again is a measure of use of hospital resources.

### **Admission duration**

Mean length of stay at our centre was increased overall for the control group, 4.0 days compared with 3.1 days in the immediate dilatation group, albeit not statistically significant ( $p = 0.19$  using Student T Test).

Despite the increased complexity of the initial procedure in the immediate dilatation group, mean length of stay for the initial procedure was unchanged (2.2 days compared with 2.0 days in the control group –  $p = 0.60$ )

Although the length of stay across admissions did not significantly alter between groups, the median time between first and second procedures in the control group was 50 days. This means that patients receiving our previous standard care (the control arm) experience a substantially longer treatment journey and have a longer illness experience, compared with those undergoing the immediate dilatation approach.

### **Procedural time**

Procedural duration was longer in the initial procedure for the more involved, immediate dilatation procedure (mean 35.9 minutes vs 26.9 minutes for the control group –  $p = 0.008$  using Student T-Test).

However, total procedural time, (combining duration of all procedures) was longer in the control group (mean 70.3 minutes compared 52.0 minutes for the immediate dilatation group) –  $p = 0.038$  using Student T-test.

### **Anaesthetic time**

Anaesthetic time was used as a surrogate for total time in the endoscopy suite and therefore resource use for the endoscopy suite itself.

Overall anaesthetic time (combining duration of all anaesthetics) was not significantly improved in the Immediate Dilatation arm (mean 82.1 minutes compared to 105.8 minutes in the control group –  $p = 0.07$  using Student T-test).

Similar to the procedural time, anaesthetic time in the index procedure was significantly longer in the Immediate Dilatation group (mean 54.0 minutes compared to 41.8 minutes in the control group –  $p = 0.03$ ).



## Safety

### Complication rate

There was no difference in overall number of complications between the two groups. 4/26 (15%) of participants in each arm suffered a complication during their overall journey ( $p = 1$  using Student T-test)).

**Figure 24: Complication Summary**

Complication	Immediate dilatation	Control
Pancreatitis	1	3
Bleed	1	0
Cholangitis	2	0
Perforation	0	1
Mortality	0	0
Total	4	4

As complication rate is generally described per procedure in the literature, complication rate per procedure was calculated for benchmarking purposes – a complication occurred during approximately 9% of our ERCP procedures. Full details of all complications including their grading using both Cotton criteria and Comprehensive Complication Index is found in Appendix E.

### Pancreatitis

As expected, pancreatitis was the most common complication and occurred in four participants; one in the immediate dilatation group, three in the control group. This was not statistically significantly different.

Three of these occurred during the first procedure. These were all mild and had all received rectal indomethacin as prophylaxis during their procedure.

One participant in the control arm incurred moderately severe pancreatitis during their second (delayed dilatation) procedure. They had no risk factors for post-ERCP pancreatitis and hence had not received indomethacin prophylaxis at their second procedure as per our institutional protocol. This participant required a stay of seven days, but required no intervention and had no organ failure.

### **Bleeding**

A single participant in the immediate dilatation group had a significant bleed following their *second* procedure whereupon they were dilated to 20mm.

This participant represented eight days after this procedure with malaena and a haemoglobin drop. A repeat duodenoscopy did not show any active bleeding and no intervention was required. Given this complication occurred during the second procedure, this is more likely a complication of the extent of dilatation, rather than a complication of the immediate dilatation approach itself.

### **Perforation**

Duodenal perforation occurred in a single elderly participant in the control arm. During her second ERCP for delayed dilatation, the duodenum was perforated by the duodenoscope itself while negotiating the duodenum prior to cannulation. This was not a complication of biliary dilatation itself. The perforation was recognised intra-operatively and was repaired immediately via an open surgical approach, along with surgical bile duct exploration and duct clearance. This patient required a short period in ICU post-operatively, but returned home without further complication.

### **Cholangitis**

Post-ERCP sepsis occurred in two participants, both within the immediate dilatation group. The first patient had mild cholangitis, with fevers only and required several days of intravenous antibiotics in a peripheral hospital following their procedure.

The second patient developed cholangitis after their *second* procedure. In this case a stent had been placed due to incomplete biliary clearance and this unfortunately fell out three days post-ERCP, requiring both antibiotics and repeat ERCP to replace the stent.

### **Comprehensive Complication Index**

Comprehensive Complication Index was calculated for each participant, as a means of stratifying and comparing the diverse complications by their severity and impact on patient care. Full details for individual patients are calculated in Appendix E. Mean Comprehensive Complication Index was similar in both groups: 3.82 in the immediate dilatation group (standard deviation 8.14) compared to 3.44 in the control group (standard deviation 9.3) ( $p = 0.874$  using Student T-test).

To contextualise this number, our study's mean Comprehensive Complication Index of 3.63 can be compared to that of Borchert's randomised controlled trial of elective cholecystectomy. They showed a mean Comprehensive Complication Index of 3.3 for straightforward elective laparoscopic cholecystectomy for biliary colic (42). This figure demonstrates the rate & severity of complications for our subset of complex ERCP to be similarly safe as the most common routine laparoscopic general surgical operation.

### Radiation exposure

Exposure to ionising radiation is an integral part of imaging the bile duct for ERCP. As seen in the below table (Figure 25), there was an extremely wide range of both radiation dose and time exposed to radiation seen in this trial, with a number of extreme outliers.

**Figure 25: Radiation dose**

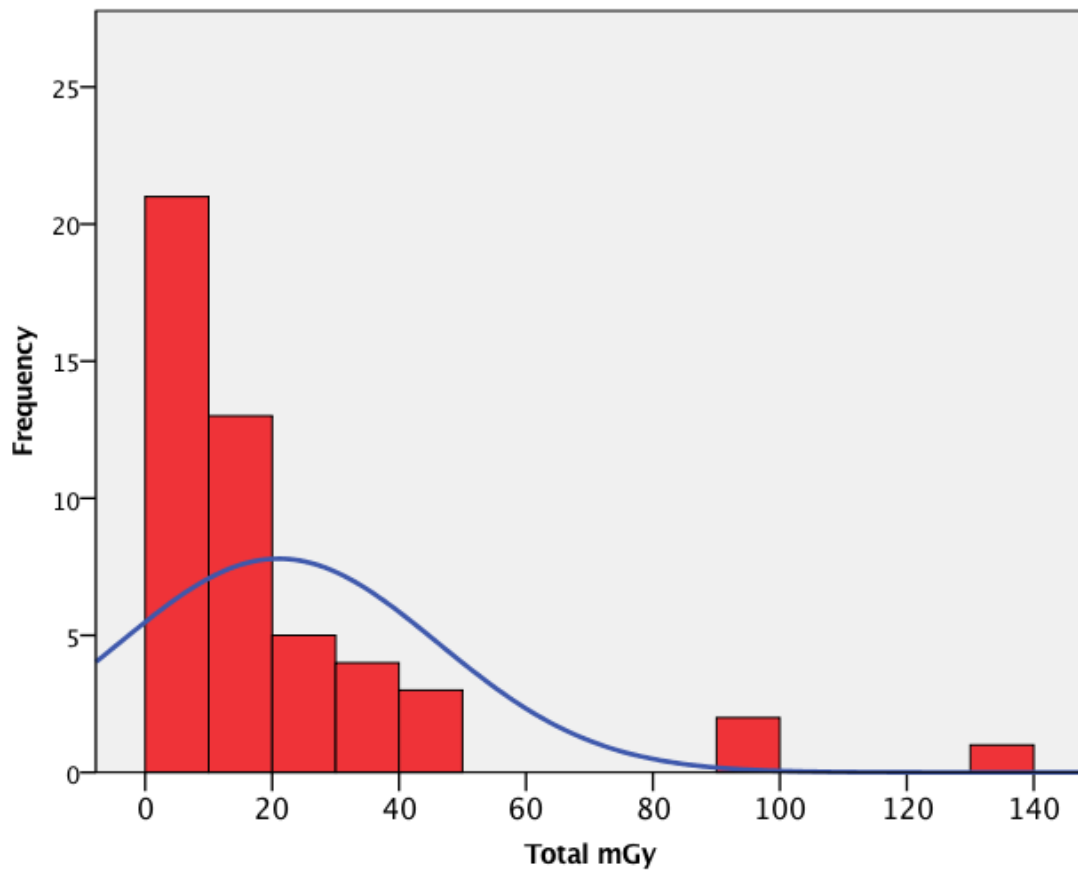
<b>Radiation Dose</b>	<b>Immediate Dilatation</b>	<b>Control</b>
<b>Median (mGy)</b>	9.5	17.5
<b>Range (mGy)</b>	0.5-130.5	2.69-47.95
		$p = 0.032^d$

<b>Radiation duration</b>		
<b>Duration – median (secs)</b>	49	97
<b>Duration range (secs)</b>	19-342	19-328
		$p = 0.032^d$

d - Mood's median test

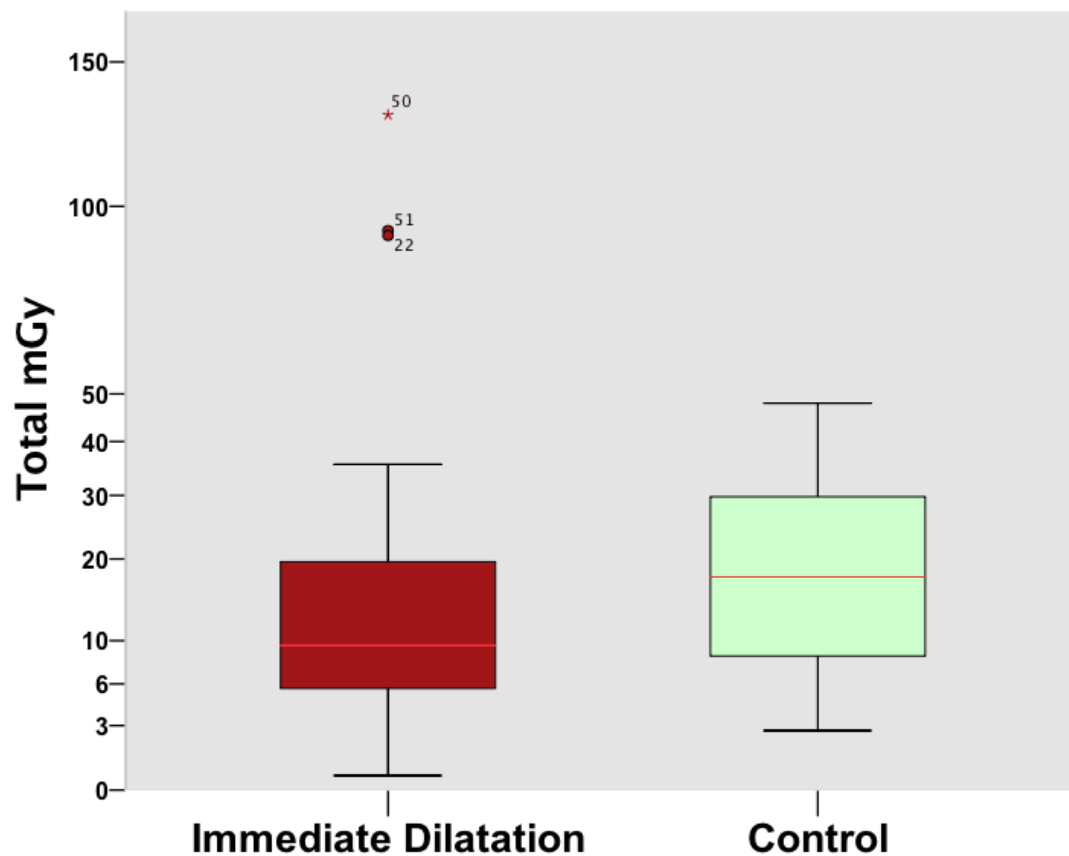
Radiation exposure shows a wide variation in radiation dose (0.5 – 130.5 mGy). As seen in Figure 26, radiation dose follows a non-parametric distribution, heavily skewed with a number of extreme high-dose outliers. Given these outliers, Mood's median test was used to compare the two arms. Skewness and kurtosis are calculable, using Z values given the relatively small sample size: Skewness 2.7 (z-value 8.0), kurtosis 8.5, (Z-value 12.8), which confirm the non-parametric distribution.

**Figure 26: Radiation dose histogram**



Absolute radiation dose was halved by using combined sphincterotomy + immediate dilatation approach (median 9.5mGy) compared with the control group (median 17.5 mGy) ( $p = 0.032$  using Mood's Median test) (see Figure 27). This equates to saving each participant in the intervention group the radiation equivalent to 5.4 years of background radiation in Australia(43).

**Figure 27: Radiation dose**



Overall duration of exposure to ionising radiation was also compared and was significantly decreased in the intervention group (median 49 seconds) when compared to the control group (median 97 seconds) –  $p = 0.032$  using Mood's Median test. Once again, there was a wide range of durations in both groups (19 - 342 seconds radiation duration).

## Discussion

### Introduction

As previously discussed, there are multiple potential approaches to dealing with stones in the bile duct. As clinicians, our decision as to the most appropriate approach comes down to a balance between both procedural efficacy and safety for our patient. In comparing the two variations of ERCP in this trial, both efficacy and safety must be examined in turn.

Firstly we discuss our procedural performance with relation to cannulation rates and complex sphincterotomy techniques. This allows benchmarking of our performance against results of seen in large multicentre series.

Efficacy has been measured in a number of ways and the clinical application of each of these is compared and contrasted, both for the arms of the study and benchmarking to the literature. Multiple ERCPs may be required for a number of reasons, not all of which are related to duct clearance. These reasons are explored, describing the various roles ERCP performs during the entire patient experience of biliary clearance. Secondary measures of efficacy, such as procedural and anaesthetic time are explored.

Finally, safety is compared both between arms and to appropriate benchmarks in the literature. Safety will be examined with regards to overall complication rate and well as rates of the individual major complications: bleeding, cholangitis, pancreatitis and perforation. The Comprehensive Complication Index is explored as a more nuanced means of comparing complications of varying severities between the groups. Finally, the ERCP-related harm of radiation is compared, discussing the clinical importance of the improvement in ERCP-related radiation exposure seen by using the immediate dilatation approach.

### Procedural performance

ERCP is a highly technical endoscopic procedure, with adequate experience and technical proficiency key to good performance. Within this already difficult procedure, extraction of large stones represents a subset of ERCP with increased complexity. A number of metrics were examined to benchmark our performance, ensuring our technical performance of ERCP in general is appropriate, independent of the stone extraction techniques of the trial.

140 participants were originally consented for this trial, with successful bile duct cannulation occurring in 132 (94%). This compares favourably to international benchmarks e.g. Kapral *et al's* study attempting to benchmark ERCP performance in Austria,

demonstrating a 86% cannulation rate for ERCPs performed by high-volume (>50 ERCPs/year) endoscopists(39). This confirms appropriate technical performance of ERCP overall within our study, particularly considering that almost all procedures had trainee involvement.

A second metric for technical proficiency commonly quoted is the use of needleknife or precut techniques for sphincterotomy and biliary access. This was 23% in our study – significantly higher than that quoted in most series. By comparison the high volume centres in the Gallriks database of all Swedish ERCP quote a 17.4% use of these more technically challenging techniques once patients who have already had previous sphincterotomy are excluded(6). Our unit has a policy of early use of precut techniques in cases of difficult cannulation. This approach has been shown to be associated with a lower risk of complications compared with prolonged attempts with standard techniques(41). Our higher rate of needle-knife or precut techniques indicates a higher proportion of difficult cannulations and our deliberate attempt to minimise the risk of post-ERCP pancreatitis.

Thirdly, complication rate may be benchmarked against that of the literature to confirm our overall performance to be safe. Once again, the patients in this trial represent a more technically challenging and therefor potentially hazardous subset of procedures than the average ERCP. Despite this, our complication rate of 9% per procedure compares appropriately to the 10% risk of complications seen in the Gallriks registry(6).

## **Efficacy**

### **Choosing an appropriate endpoint**

Efficacy of ERCP can be measured in multiple ways, to compare our study's two arms. For our patient population, larger stones have traditionally required multiple endoscopic procedures, if not open surgery. Historically, around the world (and at our institution), patients fitting within inclusion criteria of this study would undergo at least two ERCPs using the standard practice. For a patient, being potentially able to clear the duct in a single procedure is an obvious benefit, even if this is unable to be achieved 100% of the time. This makes the number of procedures undertaken for each participant the most obvious marker of the efficacy of each technique.

A number of other endpoints are commonly discussed in the literature, but all have potential issues within our trial. The most commonly described measure of efficacy in the relevant literature is duct clearance at the first ERCP. This is problematic for our study where one arm deliberately avoids stone extraction during the first ERCP, meaning that in



the control arm, clearance rate at first ERCP will, by definition, be 0, given that we are dealing with large stones, and it cannot be said that duct clearance at the first procedure is the norm. Rather, most would previously have required multiple procedures, making this measure less relevant.

One can compare extraction rates after the first procedure in which attempted stone extraction occurs. That is, successful clearance at the *first* ERCP in the immediate dilatation arm compared with that at the *second* ERCP in the control arm. From a patient perspective, it seems unlikely that they would consider it equivalently efficacious having two procedures compared to one. We have, nevertheless, compared this extraction rate at first attempt to benchmark our performance against that seen in the literature.

The above two measures of duct clearance are generally described in relation to radiological clearance on final cholangiogram at the completion of the procedure. However, this measure is flawed, as a number of patients will require further ERCP procedures despite radiological evidence of clearance at their first ERCP. Neither patients nor proceduralists would generally consider this a success. The measure we have chosen of total number of ERCPs required better captures the whole patient experience.

At the other end of the spectrum, the final commonly reported measure of efficacy is overall extraction rate. This compares the ability of the approach in each arm to extract all stones, given an unlimited number of attempts. Again, we have documented this for benchmarking purposes, but it is unlikely to be particularly discriminatory between our arms as eventual success is likely to be similar, given enough attempts.

### **Procedures to clearance**

Our primary endpoint was the total number of ERCP procedures performed on each participant. This most closely represents the patient's experience of biliary clearance and demonstrates the degree of the advantage in using an immediate dilatation approach. As previously described, for participants in the immediate dilatation group, 17/26 participants (65.4%) required only one procedure overall. 23/26 participants achieved bile duct clearance with two or less procedures, compared to 21/26 dealt with in the minimum two procedures in the control arm (see Figure 28).

**Figure 28: Total Number of ERCPs**

		No. procedures to clearance				Total
		1	2	3	4	
Arm	Immediate Dilatation	17	6	1	2	26
	Control	0	21	5	0	26
Total		17	27	6	2	52

Chi-square analysis comparing the number of ERCPs required in each arm was performed. This confirms the statistical significance ( $p < 0.001$ ) of this clinically significant improvement in care by an immediate dilatation approach.

Until the development of the combined sphincterotomy/large-balloon sphincteroplasty technique, many of these patients would have required a laparotomy and operative bile duct exploration. Even with the advent of this combined technique, our previous standard practice (the control arm) would have mandated at least two ERCPs for these patients. The ability to offer two-thirds of patients clearance with a single ERCP is a substantial improvement in clinical practice. Clearance rate after a single ERCP was compared between arms using Fisher's exact test and again this confirms statistical as well as clinical superiority ( $p < 0.001$ )

Nevertheless, a third of patients still require more than one ERCP. There are a number of reasons why further ERCP might be necessary in either group and these do not all relate to clearance of the bile duct. Further ERCPs could also relate to the need for stent removal, or to issues at subsequent cholecystectomy, such as stones dropped into the bile duct or treatment of bile leak. These issues shall be dealt with in turn.

### **Barriers to endoscopic duct clearance**

Stone extraction does not always succeed at first attempt. This may be due to very large stones, multiple stones, anatomical barriers to extraction such as a tapered or angulated distal duct or stones lodged in awkward positions within the ducts.

Very large stones (>15mm) are not expected to be able to be removed at the index procedure in our trial. Our dilatation protocol limited the initial dilatation to 15mm only, due to safety concerns. As such, the small number of patients with very large stones was always going to require a second procedure even in the immediate dilatation arm. On occasion, placement of the stent for a patient in the control can cause stone fragmentation and erosion, meaning that a 15mm dilatation may become sufficient by the second procedure.

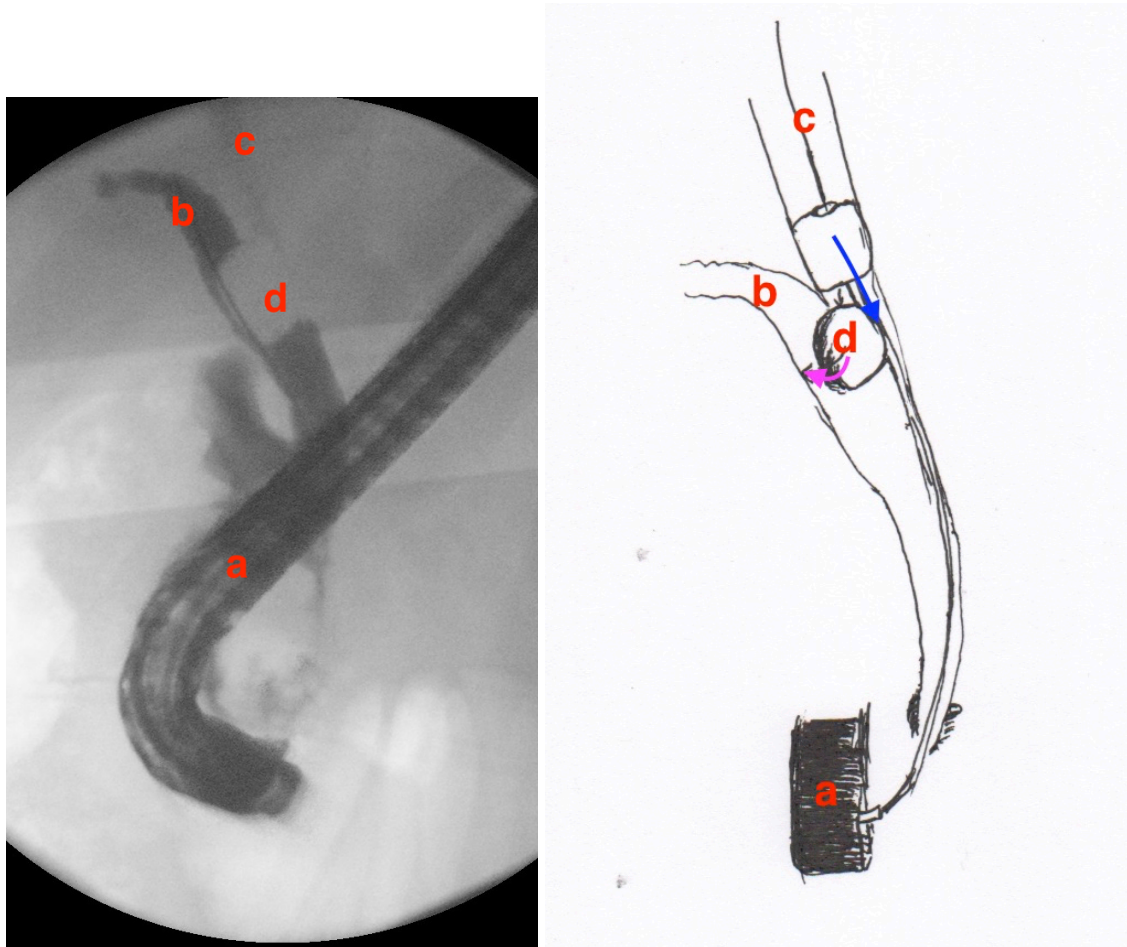
Multiple stones are another well-described factor predicting failure of bile duct clearance. As the number of stones within the bile duct increases, the chance of “missing” a stone when extracting increases. Concern about these residual stones is a reason one may leave a stent after the initial extraction attempt, which necessitates a further ERCP. By contrast, a delayed-dilatation approach deliberately leaves a stent. This stent placement may cause stone fragmentation and erosion, increasing the ability to clear at the next ERCP. Indeed Horiuchi’s study of 40 patients with large or multiple stones showed the simple intervention of placing a stent and waiting two months showed a median decrease in stone number from four to two (27). This suggests that an immediate dilatation approach may offer less of an advantage in this small subset of patients.

A number of anatomical barriers may make stone extraction fail at first attempt. A stone trapped above a narrow or angulated distal bile duct is anatomically difficult to remove. One of the benefits of a combined sphincterotomy + sphincteroplasty approach is that it allows enlargement of the entire distal duct. Nevertheless, persisting post-dilatation stricture or angulation still makes stone extraction difficult, and this difficulty would apply to both arms of our trial.

A particular anatomical issue that limits stone extraction is the stone stuck in the junction between the cystic and common bile ducts (see Figure 29). One participant in each arm had this configuration. The first, in the immediate dilatation arm, who had already previously had cholecystectomy, underwent three unsuccessful attempts to extract this stone endoscopically, each of which pushed the stone across into the cystic duct. Eventually this patient was referred for endoscopic cholangioscopy and electrohydraulic lithotripsy under direct vision. This is a sub-specialist procedure, requiring specialised equipment, and is not generally available. It is certainly not a routine alternative to the use of ERCP for

larger bile duct stones as was the practice in this study. A second patient, in the control arm, underwent two ERCPs without successful extraction. Given his gallbladder was still in situ, he was booked for a cholecystectomy with concurrent operative bile duct exploration.

**Figure 29: Stone caught in junction between cystic and bile ducts**



**a: duodenoscope, b: cystic duct, c: hepatic duct, d: stone stuck in junction between cystic and hepatic ducts**

**In the figure on the right, it can be seen as the extraction balloon in the hepatic duct is pulled down (blue arrow), it can slide past the stone in the widened cystic duct junction, pushing the stone across into the cystic duct (magenta arrow) and failing to pull it out.**

As seen in these cases, stones caught in unusual positions may be a reason that ERCP extraction fails. This is not because of the immediate dilatation approach itself but rather the innate inability of a non-directed balloon to capture the stone in this position.

### **Confounders: Cholecystectomy**

Setting aside genuine reasons for stone extraction to fail, the major confounder to our primary endpoint was further procedures related to post-ERCP removal of the gallbladder (cholecystectomy). Once the bile duct is cleared endoscopically, in most cases the gallbladder (if still present) is removed to prevent further complications of stone disease.

Whilst the gallbladder remains in situ, stones may fall again from the gallbladder into the bile duct, either whilst waiting for cholecystectomy, or during manipulation of the gallbladder intra-operatively. In Pierce *et al.*'s series, further bile duct stones were seen in as many as 21% of patients having completion cholecystectomy after ERCP(44), stones which may cause further complications in the intervening time. This number obviously increases if the time to cholecystectomy increases.

For this reason, some endoscopists prefer to routinely leave a prophylactic biliary stent even if the bile duct is cleared in patients who are planned to have a cholecystectomy. This prevents complications from recurrent stones before or during cholecystectomy, albeit at the cost of a further procedure to remove the stent. This is a potential confounder for our primary endpoint as it means an increase in the number of procedures required overall, even though the bile duct is clear at the end of the first procedure. Nevertheless, we have not changed the endpoint as this accurately represents the patients' experience, rather than purely measuring technical success at a single point in time as most trials do.

If a stent is not left, Pierce *et al.* suggest 21% of patients undergoing cholecystectomy will have recurrent or residual stones found in the bile duct(44). Given the risk of recurrent cholangitis and pancreatitis, it is necessary to remove these stones with repeat ERCP, and indeed this was seen in two participants who were not stented, but were found to have recurrent or residual stone at a later cholecystectomy. It is worth noting both of these patients had unexpectedly delayed their cholecystectomy for several months for personal reasons.

In addition to retained stones, ERCP may be required to deal with complications of cholecystectomy. One of our participants required two ERCPs to deal with a bile leak after cholecystectomy, despite clearance at the first procedure. Overall this means that, while one technical clearance at first procedure is high in the immediate dilatation arm, further procedures may be required nevertheless, biasing our primary endpoint in favour of the control.

Our literature review has not come across any other trial specifically reporting data in this fashion. As a surgical unit, we provide both endoscopic bile duct clearance and

cholecystectomy, including surgical clearance of the bile duct where necessary. This represents the full process of treatment of gallstones for our patients.

By comparison, the majority of published trials (if they actually specify their methods of determining success) report on % of participants who have apparent clearance at the end of their first procedure. For the reasons explained above, this measure will overestimate the effectiveness of ERCP in this patient population, ignoring the group who will pass further stones down the bile duct prior to cholecystectomy, or who will need ERCP for a complication of surgery. Whilst this may be a technical success, most patients would not believe the procedure successful if they still need to have it repeated later on. As a surgical unit, rather than endoscopy service alone, we have been able to capture the full patient experience, which is a strength of our study design.

### **Secondary endpoints of efficacy**

Most trials describing efficacy of ERCP do not capture the full patient experience in the way we have attempted to in this trial. Two other endpoints are more commonly used in these studies to describe efficacy – % of bile ducts cleared at first attempt and overall clearance rates.

#### **Clearance at first attempt**

Clearance at first attempt is generally defined (if specified) as no stones seen on cholangiogram at the completion of the procedure. With regards to the immediate dilatation group, this radiological success was achieved in 88% of participants, which is very similar to the rates reported in Jin's meta-analysis of randomised controlled trials involving large stones (82.2%). The similarity of clearance rates benchmarks our study, demonstrating at least equivalent proficiency with the technique itself to that reported in comparable randomised controlled trials.

Comparing arms within our trial, we can compare the % clearance after procedure one in the immediate dilatation arm to the clearance after *procedure two* in the control group. This compares success of our *first attempt* at duct clearance in each arm, as control patients receive sphincterotomy and drainage alone at their first procedure, with stones left in situ. There was no significant difference between this measure of clearance at first attempt between the two arms: 88% in the combined approach vs 84% in the control group.

We had previously hypothesised that delaying the dilatation to allow sphincter remodelling may improve the efficacy of dilatation. In addition, the presence of a stent has previously been shown to improve efficacy of removal of large or multiple biliary stones(27), presumably by the mechanism of the stent eroding and fragmenting the stones. If either of these effects were present, they were not large enough to generate a difference between our two groups.

### **Overall clearance**

By the conclusion of all their procedures, all patients in the immediate dilatation group had their duct cleared by ERCP techniques, albeit one required referral to a secondary centre for endoscopic cholangioscopy and electrohydraulic lithotripsy (3.8%). This patient had a stone lodged in the cystic duct/hepatic duct junction. By contrast, three participants in the control group (11%) failed endoscopic clearance and required surgical bile duct exploration ( $p = 0.24$  using Fisher's exact test).

We see that overall clearance is not significantly different in either arm, given enough attempts. This reflects the fact that overall failure of endoscopic clearance (for conventional ERCP) reflects the size of stones and anatomical barriers to clearance, and the trial intervention (immediate rather than delayed sphincteroplasty) does not influence this variable. For benchmarking purposes, our overall clearance rate compares favourably with that reported in the major meta-analyses, again demonstrating adequate technical performance in our trial(15).

### **Indirect measures of efficiency**

We have reported on a series of indirect measures of efficacy in this trial. Procedural and anaesthetic time and length of stay are measures of efficiency, in particular relating to resource allocation.

### ***Procedural time and anaesthetic time***

We have considered the two together here, as the data are similar. From a practical point of view, procedural time is the time the endoscopist is actually performing the procedure as measured by routine nursing documentation independent from this trial. This is the most commonly reported measure of time. Anaesthetic time considers the entire time the patient is within the endoscopy room, again measured by routine nursing documentation. There

are two relevant measures here: time for the initial procedure and overall time. Both procedural and anaesthetic times are proxies for efficiency of usage of the endoscopy suite.

Across the trial, overall procedural time and overall anaesthetic time were both lower in the immediate dilatation arm. This reflects the additional time cost of the additional procedures in the control arm. Thus overall, a combined approach is more resource-efficient overall.

However, it must be conceded that the first procedure and anaesthetic were both significantly longer in the intervention group. Most ERCPs are performed on a semi-urgent basis. Routine immediate dilatation would potentially slow down these lists, which is relevant in terms of planning theatre usage on a day-to-day basis. However, fewer follow-up procedures would be required, improving theatre utilisation overall.

### ***Length of stay***

Length of stay is an easily measurable indicator of efficiency of usage of hospital resources. Unlike the above other measures of time, total length of stay was not markedly different between the two groups, with a median total length of stay of three days in each group.

In general, the major determinant of length of stay for patients undergoing ERCP relates to their initial, acute or semi-acute presentation. Hence most of the length of stay relates to the treatment of jaundice, cholangitis or pancreatitis (if present), rather than to the ERCP. Subsequent additional procedures were predominantly straightforward day-case procedures (unless a complication occurred). In addition, a series of logistic issues (e.g. arranging transportation for transfer, social supports and discharge planning) have a substantial impact on length of stay, particularly during the index admission, meaning the benefits for a single-stage procedure with regards to length of stay are relatively modest.



## **Safety**

It seems clear that for clinically relevant measures, an immediate dilatation approach is more efficacious than the staged approach seen in our control arm. We now compare the relative safety of the two approaches, comparing raw numbers of complications, each major group of complications in turn and the role of the Comprehensive Complication Index. In addition, radiation exposure is a potential harm integral to any ERCP procedure. The overall radiation dose per patient is small during most procedures, but the cumulative dose to endoscopy staff is a substantial concern(45)

## **Complications**

Overall, complications occurred equally across each arm of our trial – 15% of participants suffered a complication during any of their multiple procedures. No participant suffered more than one complication overall.

With regards to individual procedures, a non-statistically significant decrease in complication rate *per-procedure* was seen in the control group. The increased number of procedures required in this arm however balanced this identically.

What this means is the risk to participants is identical across the arms; all the risk is concentrated in one procedure with an immediate dilatation approach, as opposed to spreading the same risk across two procedures as one moves the dilatation to a second procedure. From a patient perspective, it seems unlikely that any patient would choose to undergo a second procedure just to spread the same total risk across the two procedures.

## ***Comparison to literature***

Generally, ERCP complications are reported per procedure, rather than per participant as reported in this study. For benchmarking purposes this per-procedure figure was calculated to compare our complication rates to the literature.

Overall, 9% of procedures incurred a complication. This benchmarks appropriately with the largest collection of randomised data regarding sphincterotomy and sphincteroplasty: 9.9 – 10.5% (10). This suggests our performance of this procedure is appropriately safe.

Our rate is, however, higher than the 5% reported in Meine's meta-analysis of available studies of the combined technique. There are a number of reasons for this, which speaks largely to the carefully curated, non-randomised group of cases described in Meine

and, indeed most case series on this topic. We have deliberately tried to keep our inclusion criteria as close to wider clinical practice as possible, so higher risk cases were not excluded, as they often are in previous series. Patients in whom precut cannulation techniques were required were included, as were patients with multiple pancreatic duct cannulations or injection of contrast into the pancreatic duct – all known risk factors for complications, particularly pancreatitis(3). In addition, as a teaching hospital, all procedures had involvement of an endoscopist undergoing advanced training in ERCP for a major portion of the procedure, which are often excluded.

Finally, our follow-up protocol included either clinic review or follow-up phone-call for all participants, as is part of our routine practice. This allows comprehensive and accurate capturing of all complication data. This is in contrast to most studies, which do not specify any follow-up beyond the index admission and may miss delayed presentations of pancreatitis or bleeding after discharge – 38% of complications in our study occurred after discharge from our centre and would have been unreported without our follow-up protocol. Indeed our complication rate would have been just 5.6% should these delayed complications have been missed with less rigorous follow-up – identical to the rate in Meine's meta-analysis.

### **Pancreatitis**

Pancreatitis is the most common complication following ERCP(46) and this finding was reproduced within our study. Four participants overall suffered post-ERCP pancreatitis. No participants developed pancreatic necrosis, or required surgical intervention or intensive care admission.

Pancreatitis is more common in the initial procedure, as cannulation of a “virgin” ampulla requires considerably more manipulation. Three of the four cases of post-ERCP pancreatitis occurred in the first procedure in our trial.

Several risk factors for post-ERCP pancreatitis have been defined. These include younger age, need for precut sphincterotomy techniques, previous history of post-ERCP pancreatitis, prolonged attempts at cannulation, pancreatic duct injection and normal bilirubin(3). We did not find a correlation with these factors and pancreatitis within our trial.

At our institutions (and within this trial) we practice a policy of administration of rectal indomethacin 100mg during the procedure for all but the lowest risk procedures. This means this is administered at every initial ERCP, and at subsequent ERCPs if any of the

above risk factors were present. Multiple trials support the use of rectal anti-inflammatories in decreasing the risk of post-ERCP pancreatitis, particularly in high-risk populations(47). The evidence regarding use of rectal anti-inflammatories in lower risk procedures is conflicting(48), with a recent meta-analysis supporting our policy of restrictive administration(36). We did not see a relationship between use of rectal indomethacin and post-ERCP pancreatitis in our trial.

### **Bleeding**

A concern had been raised in our previous retrospective review that an immediate dilatation approach may cause a higher rate of clinically significant bleeding than the staged approach in our control arm(29). Our previous data had suggested a three-fold increase in bleeding using an immediate dilatation approach.

Only one participant in this trial developed ERCP-related bleeding. This participant was within the immediate dilatation arm, however the bleed occurred following their *second* procedure at which a second dilatation (to 20mm) was deemed necessary. Given it occurred following the second procedure, this bleed should be considered a complication of the extent of dilatation, rather than a complication of the immediate dilatation approach itself. This is in keeping with previously published data suggesting dilatation >15mm may be associated with an increased risk of serious complications(49).

### **Perforation**

Perforation is thankfully a rare complication. It may occur due to any of three mechanisms. Firstly, it may occur from the mechanical trauma of the duodenoscope itself negotiating the duodenum, prior to any biliary intervention. Secondly, on cannulation, the wire can pass out through the wall of the duodenum and into the retroperitoneum. This, if recognised, can usually be managed conservatively. Finally, the act of dividing the biliary sphincter (sphincterotomy) or indeed stretching it with the dilating balloon can rupture the duodenum or the lower part of the bile duct. This final mechanism is one major concern with performing an immediate dilatation approach, but thankfully not a complication seen during our trial.

There was a single duodenal perforation, which occurred in the control arm. This patient was in the delayed sphincteroplasty group and was undergoing her second ERCP. The duodenum was perforated by the duodenoscope itself while negotiating the duodenum prior to cannulation. This was not a complication of biliary dilatation itself. Nevertheless, it

may not have occurred if this participant had her bile duct cleared during her first procedure using an immediate dilatation approach.

The very low incidence of this complication limits abilities to directly compare perforation rate in any comparative trial. Significant differences between approaches of this rare complication can probably only be identified by retrospective review of multicentre audits.

### **Cholangitis**

Post-ERCP sepsis occurred in two participants, both within the immediate dilatation group. The first patient had mild cholangitis after her first procedure, which resolved with intravenous antibiotics. The second patient developed cholangitis after their *second* procedure. In this case a stent had been placed due to incomplete biliary clearance and this had unfortunately migrated out of the bile duct, leading to cholangitis.

Both of these cases speak to the importance of adequate biliary drainage in preventing biliary sepsis. Once the bile duct has been cannulated, bacteria from the duodenum enter the previously sterile confines of the biliary tree. If the bile duct is draining freely through a widely patent ampulla, this rarely causes a problem. However if there is residual stone or sludge within the bile duct, drainage may be impaired and bacteria can multiply, causing cholangitis. Stent placement can prevent this, albeit at the cost of the need for a further ERCP to remove the stent. Hence cholangitis would be expected to be lower in the control arm by design. At our centre, we follow a policy of leaving a stent if there is any concern about either the adequacy of our bile duct clearance or presence of significant infection at the time of the ERCP. This policy may bias the primary endpoint of our trial against the immediate dilatation arm (raising the number of ERCPs), but seems a safer way to proceed.

### **Comprehensive Complication Index**

There are a number of problems with the above measures of simple complication rate and rates of each individual complication. Our overall complication rate lumps together complications as dissimilar as mild pancreatitis requiring no treatment but observation for 72 hours, and a duodenal perforation requiring a laparotomy and intensive care. Comparing rates of the individual complications allows a fairer comparison. However some complications e.g. perforation are so uncommon as to make effective comparison practically impossible.

The Comprehensive Complication Index was developed as a more nuanced means of comparing multiple potential complications of surgical procedures. It allots a score weighted by severity for each complication during a patient's admission, and then adds these together to provide a composite score from 0 (best) to 100 (worst). This composite score combines both number and severity of all complications across a patient's admission and shows high correlation with both post-operative health status and patient ratings of complications(32). By summarising all post-operative complications and their severity, it allows a more sensitive measure of comparing surgical complications (particularly low-incidence ones) than simple complication rate.

Whilst it has not specifically been used in ERCP literature, the Comprehensive Complication Index has been validated against multiple surgical procedures (32, 33), showing improved discrimination and ability to compare distinct complications in a meaningful manner. It has now been used in over 50 published trials as a primary measure of complications and is becoming accepted as a more appropriate measure of complications in the surgical literature.

A strength of the Comprehensive Complication Index is that it accounts for multiple complications occurring to the same patient. Thankfully, in our study, no patient suffered more than one complication.

We saw no significant difference in mean Comprehensive Complication Index between the two arms of our trial (mean 3.82, standard deviation 8.14 in the immediate dilatation group compared to mean 3.44, standard deviation 9.3 in the control group ( $p = 0.874$  using Student T-test)). The lack of difference seen here even using this more sensitive tool provides additional support to the similarity of the two arms with regards to complications.

As previously stated, our study's mean Comprehensive Complication Index of 3.63 cannot be compared to other endoscopic studies, as this is (to our knowledge) the first use of this metric in describing complications for an endoscopic procedure. It can however be compared to other biliary procedures, such as Borchert's randomised controlled trial of elective cholecystectomy. They showed a mean Comprehensive Complication Index of 3.3 for straightforward elective laparoscopic cholecystectomy for biliary colic(42), very similar to our number for complex ERCP.

A comparable tertiary Australian centre recently published their results of acute cholecystectomy in patients with bile duct stones found on cholangiogram(50). These patients either had ERCP or laparoscopic bile duct exploration together with cholecystectomy. Across their whole patient population, they described an overall mean

Comprehensive Complication Index of 6.88 (SD 11.37). While the addition of cholecystectomy would be expected to increase the figure over ERCP alone, our mean figure of 3.63 (SD 8.65) across our trial therefor benchmarks favourably with that of cholecystectomy and ERCP in a comparable patient population, suggesting our approach is appropriately safe.

### **Limitations**

We can conclude from our complication data that complication rates are essentially identical across both groups. Whilst the number of complications per procedure is slightly higher in the immediate dilatation group, this is offset by the increased number of procedures in the control group, meaning there appears to be no safety benefit to delaying the procedure. In essence, the risk is brought forward in the immediate dilatation group, but the overall risk is unchanged.

There are some limitations to conclusions with regards to complications, however. Chief amongst these is the difficulty of comparing relatively rare complications in a fairly small randomised trial. This means that while, at face-value, there does not appear to be a clinically significant difference, it is true that the trial is underpowered (even without considering a Bonferroni correction) for statistical analysis.

Using the Comprehensive Complication Index offers the potential to discriminate better between groups, by combining several low-incidence complications and hence improving the sensitivity of the analysis.

Sample size calculation using the Comprehensive Complication Index data seen in our trial suggests that, assuming the proportions in our trial to be truly representative, using an  $\alpha$  of 0.05 and a  $\beta$  of 0.2, 13082 participants would be required to show a statistically significant difference between arms. Even for this more sensitive metric, the conclusion is that it would be impossible, in practical terms, to demonstrate statistical significance for the small (and clinically insignificant) differences observed between groups.

### ***Radiation dose***

There is a final measure of potential complication intrinsic to every ERCP – the exposure to ionising radiation. Prolonged ERCP for difficult biliary stones is one of the most radiation intensive procedures in general surgery. Radiation exposure leads to risks to both patient and staff alike, in particular to the endoscopy nursing staff who may be exposed to radiation from this source on a daily basis.

As previously discussed, overall duration of radiation exposure and absolute radiation dose were both halved by using an immediate dilatation rather than a staged approach. This is the equivalent of saving each patient the 5.4 years of Australian background radiation(43). The issue is that of increasing the patients' long-term risk of developing malignancy. This cancer risk depends on both radiation dose and the patient's age, with a 20-year-old female, on average increasing lifetime risk of malignancy by 0.2% by the additional radiation of the combined approach in the control group(51).

By contrast, for the average age of our patient population, an immediate dilatation approach decreases procedure-related cancer risk only by 0.04%, with cancer risk ascribed to the procedure decreasing as patients' age increases. It is worth noting that lifetime cancer risk across our community is between 37 - 45%, meaning that even for our theoretical 20-year-old female, cancer risk would only increase from 37% to 37.2%(52).

In addition to the patients themselves, endoscopy staff are also exposed to ionising radiation from ERCP and it is they who probably derive greater benefit from minimising radiation exposure. Estimates of occupational radiation exposure for endoscopy staff vary widely, with estimates of annual safe limits for the primary proceduralist varying from as few as 89 procedures (53) to >100,000(54) to significantly increase the risk of radiation-related cancer. In addition to the risk of cancer to endoscopy staff, more current understanding of radiation-related risks suggest risk of cataract formation in the lens of the eye at relatively low doses. This would suggest that radiation-related risks would be closer to the lower of these estimates, and may be exceeded in normal practice of a busy ERCP unit(45) if eye protection is not worn.

As can be seen in our results, radiation doses vary widely, in both arms of the trial (0.5mGy – 130.5mGy total exposure across all procedures). Whilst some of this relates to differences in dilatation approaches, a large portion of this depends on other factors. These may include difficulty of cannulation and stone extraction, need for multiple stents, patient position in relation to the cathode tube and patient size. Furthermore, radiation dose has been shown to vary significantly between proceduralists, their experience and their techniques(55-57).

Whilst the impact of the dilatation technique on radiation exposure is relatively small within each individual procedure, it still forms a part of the ALARA (as low as reasonably able) approach to minimising radiation exposure. This is probably more important to minimising overall occupational exposure to endoscopy staff and their long-term risk of cataracts and malignancy, rather than the relatively small exposure to each individual patient.

## Conclusions

Bile duct stones remain a source of significant morbidity in our population. Whilst multiple methods exist for their clearance, endoscopic sphincterotomy remains the most common method for removing stones. Nevertheless, in 15% of patients, stones are too large for removal with sphincterotomy alone.

Patients with these large stones have traditionally required surgical exploration of their bile duct, usually with major open surgery. The technique of adding large balloon dilatation to sphincterotomy has offered a further endoscopic advance to improve clearance of these difficult bile duct stones. We have pondered as to whether it is most appropriate to perform this sphincteroplasty at the same session as sphincterotomy or to return at a subsequent procedure, as has been our previous practice. At the end of the day, this decision as to the most appropriate sequence of sphincterotomy and sphincteroplasty comes down to balance between efficacy and safety.

## Efficacy

We have shown an obvious improvement in efficacy for the immediate dilatation approach, with a statistically and clinically significant decrease in the number of ERCPs required for participants in the immediate dilatation group. Whilst this may seem self-evident, our pre-trial experience saw that many patients require further ERCPs for a series of reasons unrelated to bile duct clearance, ERCPs that are generally not captured by other trials in the literature, but are a key part of the patient experience. The ability to quantify this improvement is vital to allowing meaningful comparison with complication data, making a decision balancing both efficacy and safety.

Radiological bile duct clearance is seen in 88% of our participants undergoing sphincterotomy and sphincteroplasty in the same session. Despite this, a number of participants required further ERCPs for a series of reasons that are often not captured by previous trials. We saw 35% of participants in the immediate dilatation group requiring more than one ERCP. Nevertheless, the ability to offer two-thirds of our patients bile duct clearance with a single procedure is a clear advance on our previous practice. This is especially important as a tertiary referral centre, with many patients travelling long distances from across the state.

In addition, we see improvements across a number of measures of resource allocation. In particular, the total procedural time is improved by using an immediate dilatation approach, meaning this approach allows more effective use of limited time in the



endoscopy suite. Length of stay showed a trend towards improvement, however this was not statistically significant in this trial.

## **Safety**

Whilst we can see a clear improvement in efficiency with an immediate dilatation approach, this must be balanced against the complications of each approach. We saw similar complication rates in both arms, measured by both raw complication rate and the more sensitive novel endpoint – the Comprehensive Complication Index. Our complication rates were appropriate to benchmarks from the literature.

On one measure, we saw a significant improvement in safety – that of radiation exposure. The significant decrease in median radiation dose, whilst small in the context of a single patient, is much more meaningful to endoscopy staff, exposed to hundreds of ERCPs per year. Previous papers have suggested endoscopy staff may in fact be exceeding their safe yearly limits in the course of their routine practice, thus any significant decrease is worthwhile. We have a duty to our endoscopy staff and ourselves to minimise our occupational exposure to ionising radiation, minimising the potential harm caused to our colleagues and ourselves.

## **Subpopulations that may not be appropriate/advantageous**

There may be a number of subpopulations in which an immediate dilatation approach may not be as advantageous as the overall population. Our small sample size precludes subgroup analysis, but allows identification of a few such scenarios.

## **Cholangitis**

In patients with cholangitis, the most important aim of ERCP is ensuring adequate drainage of infected bile. In such cases, a stent is often placed to ensure drainage, even when the bile duct appears cleared. Given the patient will need a further ERCP to retrieve the stent, the benefit of a single-stage approach may be less profound.

## **Need for subsequent cholecystectomy**

In patients with bile duct stones, removal of the gallbladder is usually recommended (assuming the patient is fit) to prevent future morbidity from gallstones. The most common

reason for repeat ERCP after duct clearance in our immediate dilatation group was to deal with issues found at subsequent cholecystectomy. This was to deal with either new or residual bile duct stones seen on cholangiogram, or complications such as bile leak. We would not advocate that an immediate dilatation approach is unsuitable if later cholecystectomy is planned, but would caution that further procedures may still be required. This becomes more likely when cholecystectomy is excessively delayed, as was seen in a number of our patients requiring repeat ERCP.

### **Very large stones**

Our initial dilatation was limited to 15mm diameter. This means that there was a subpopulation of very large stones (>15mm) that we would not expect to be able to extract in the initial session in our trial. Given this, the efficacy benefit of performing sphincteroplasty at the initial session may be less profound in this group. Further evaluation of up-front dilatation to >15mm (as has been described in other studies) would be required to assess if an immediate dilatation approach is appropriate in this setting.

## **Limitations**

### **Sample size**

The obvious limitation of this study is the relatively small number of participants in each arm. Whilst it is adequately powered to document the improvement in efficacy, conclusions on other endpoints must be more guarded.

With regards to efficacy, we saw no difference in clearance at first attempt (i.e. first procedure in the immediate dilatation arm but second procedure in the control arm). A small difference in favour of the control arm may be missed given our small numbers. From a practical point of view, the limited benefit of a small difference in this endpoint would be outweighed by the inconvenience of requiring a second ERCP and is for practical purposes irrelevant.

Complication rates are clinically much more important. Whilst the complication rates reported in this study are identical in each arm, the study is not powered to show a statistically significant difference between groups for this endpoint. We have used the more sensitive Comprehensive Complication Index to further explore the differences between complications in each group and again found no significant difference between arms. What we can conclude is there does not appear to be a significant benefit to delaying

sphincteroplasty from a complication point of view. We cannot categorically state that complications of each approach are equivalent however.

### **Exclusion rate**

A high exclusion rate was seen in this study, with far more potential participants approached and consented than were eventually found eligible. This was in part a deliberate strategy, to avoid missing participants as preoperative imaging, particularly ultrasound, is unreliable at demonstrating stone size and assessing eligibility prior to ERCP.

This exclusion rate could be lowered by more liberal use of magnetic resonance cholangiopancreatogram (MRCP) preoperatively. MRCP would allow more accurate measurement of stone size and therefore save approaching patients with small stones. This would however add significant cost and potentially delay ERCP while awaiting MRCP that would otherwise not alter clinical management.

### **Strengths**

There are a number of strengths of this study. Firstly, this is the only study to provide prospectively-collected, randomised data comparing an immediate to a delayed sphincteroplasty approach. It adds substantial rigor to the evidence provided by the only two retrospective series addressing this question.

Secondly, in comparison to many series in the literature describing the combined sphincterotomy/sphincteroplasty technique, our inclusion criteria are broad and practical. In particular, patients requiring needleknife or precut access are included, as are those in whom inadvertent pancreatic cannulation is performed. In addition, the vast majority of procedures were performed by an advanced trainee endoscopist, under consultant supervision. The inclusion of these higher risk real-life scenarios allows greater applicability to the practical situations in which this technique is likely to be used.

Thirdly, we performed comprehensive follow-up of all patients, either in person or by phone call at least a week later. This means that all complications were identified and managed, rather than just those occurring during the inpatient admission, which would have missed almost 40% of complications identified.

Finally, as a surgical unit, we are in the unique position to follow the entire course of treatment for bile duct stones. The need for subsequent cholecystectomy is a key part of prevention of future bile duct stones and may in itself necessitate further ERCPs. From a

patient point of view, these procedures are a necessary part of their treatment and should be included in any endpoint assessing efficacy. This is in contrast to the usually reported measure of % clearance at first procedure, which censures further ERCP procedures as long as the cholangiogram appears clear at the end of the first procedure. Inclusion of these procedures most accurately reflects the true patient experience of treatment and prevention of bile duct stones.

## **Future directions**

This trial provides evidence that, for most patients with large stones, performing sphincterotomy and sphincteroplasty in the same procedure appears more appropriate than delaying the sphincteroplasty to a second procedure. Greater numbers would be required to show that the complication rates are equivalent, however we have not shown a significant advantage to delaying sphincteroplasty on this measure.

## **Prospective audit**

Given the rarity of significant complications, and the apparent similarity seen in the arms of our trial, it is unlikely that we could design a practical trial to show a significant difference between groups. A practical solution to the difficulties of sample size is the use of prospectively collected audit data. Prospectively collecting data on all ERCPs, as we now do at Eastern Health, allows identification of trends among much greater numbers than are possible in the confines of a randomised controlled trial, as long as the same rigorous follow-up protocol is adhered to. A similar process is seen in Sweden with their GallRiks database of all cholecystectomies and ERCPs performed has allowed analysis of outcomes of almost 60,000 ERCPs(58).

## **Subpopulations**

Of particular interest are the above-mentioned subgroups in whom the benefits of an immediate-dilatation approach may be less pronounced. Again, given the small numbers involved, careful audit of these patients in particular will be the key to continue to assess whether the immediate dilatation approach is truly appropriate for them.

## **Role with increased use of ERCP-directed cholangioscopy**

Finally, as new technology develops, different techniques become available for dealing with old problems. One such new solution is ERCP-directed cholangioscopy where a very fine disposable endoscope is passed within the working channel of the duodenoscope itself and up into the bile duct, to allow directed capture of stones in difficult positions and well as directed lithotripsy and visual confirmation of bile duct clearance. This directed lithotripsy allows more straight-forward fragmentation of stones, to allow extraction through a smaller ampullary orifice, meaning the very large orifice created by combined sphincterotomy/sphincteroplasty may not be required.

Such procedures are currently limited to a handful of centres and are likely to remain much more expensive than the combined sphincterotomy + sphincteroplasty approaches. One of the benefits of sphincterotomy + sphincteroplasty is that it uses tools that are cheap and already available in any interventional endoscopy suite, meaning the procedure is readily available to all ERCPists. Given this, it seems likely that sphincterotomy followed by immediate sphincteroplasty will remain a mainstay of large biliary stone extraction, with ERCP-directed cholangioscopy reserved for failed extraction after maximal safe sphincteroplasty has been performed.

## **Appendices**

### **Appendix A: Eastern Health Antibiotic and NSAID protocol for ERCP prophylaxis** **Non-Steroidal Anti-Inflammatory Drugs:**

- Single dose PR indomethacin 100mg suppository after sedation
- Indicated for
  - all patients with intact papilla (no previous sphincterotomy)
  - patients with past history of pancreatitis
- If not contraindicated (recurrent GI ulceration, significant renal impairment, allergy, NSAID-induced asthma)

### **Antibiotics:**

- Indicated for:
  1. Patients with obstruction where drainage is incomplete
  2. Patients who are having a repeat ERCP
  3. Patients with hilar strictures or primary sclerosing cholangitis.
  4. Immunocompromised patients, patients with potential for bacterial endocarditis
- First line: Piperacillin-Tazobactam 4.5g IV
- Second line: Cefepime 2g IV
- Third line ie severe penicillin sensitivity or cephalosporin sensitivity: ciprofloxacin 400mg IV

Given the patient population involved, most patients in this study received indomethacin, but not antibiotics.

## Appendix B: Cotton criteria

Complication	Mild	Moderate	Severe
<b>Bleeding</b>	Clinical (i.e. not just endoscopic) evidence of bleeding.  Hemoglobin drop <3g, and no need for transfusion	Transfusion (4 units or less), no angiographic intervention or surgery	Transfusion 5 units or more, or intervention (angiographic or surgical)
<b>Perforation</b>	Possible, or only very slight leak of contrast, treatable by fluids and suction for 3 days or less	Any definite perforation treated medically for 4 - 10 days	Medical treatment for more than 10 days or intervention (percutaneous or surgical)
<b>Pancreatitis</b>	Clinical pancreatitis, amylase at least three times normal at more than 24 h after the procedure, requiring admission or prolongation of planned admission to 2-3 days	Pancreatitis requiring hospitalization of 4-10 days	Hospitalization for more than 10 days, or hemorrhagic pancreatitis, phlegmon or pseudocyst, or intervention (percutaneous drainage or surgical)
<b>Infection (cholangitis)</b>	>38°C  24-48 hr	Febrile or septic illness requiring more than 3 days of hospitalisation or endoscopic or percutaneous intervention	Septic shock or surgery
<b>Basket impaction</b>	Basket released spontaneously or by repeat endoscopy	Percutaneous intervention	Surgery

Any intensive care unit admission after a procedure grades the complication as severe. Other rarer complications can be graded by length of needed hospitalization. (30)



### Appendix C: Clavien-Dindo score

<b>Grade I</b>	Any deviation from the normal postoperative course without the need for pharmacological treatment or surgical, endoscopic, and radiological interventions. Allowed therapeutic regimens are: drugs as antiemetics, antipyretics, analgesics, diuretics, electrolytes, and physiotherapy. This grade also includes wound infections opened at the bedside.
<b>Grade II</b>	Requiring pharmacological treatment with drugs other than such allowed for grade I complications. Blood transfusions and total parenteral nutrition are also included.
<b>Grade III</b>	Requiring surgical, endoscopic or radiological intervention
<b>Grade IIIa</b>	Intervention not under general anesthesia
<b>Grade IIIb</b>	Intervention under general anesthesia
<b>Grade IV</b>	Life-threatening complication (including CNS complications)* requiring IC/ICU management
<b>Grade IVa</b>	Single organ dysfunction (including dialysis)
<b>Grade IVb</b>	Multiorgan dysfunction
<b>Grade V</b>	Death of a patient

\*Brain hemorrhage, ischemic stroke, subarachnoidal bleeding, but excluding transient ischemic attacks. CNS, central nervous system; IC, intermediate care; ICU, intensive care unit. (59)

## Appendix D: Sample size assumptions

### Success

These were based initially on the proportional success in our previously published series (29). The proportional success in the immediate dilatation & control arms are seen below (adapted from table 3 of Ho's paper, adding an additional procedure to those in the delayed sphincteroplasty group as each of these will have had, by definition, a previous ERCP).

No. procedures required for duct clearance	Immediate dilatation arm	Control
1	0.7	0
2	0.1	0.7
3	0.1	0.2
4+	0.1	0.1

For the purposes of sample size calculation, the proportional success of the immediate dilatation arm was estimated more conservatively for two reasons; firstly, the retrospective nature of the previous paper may bias in favor of the intervention (immediate dilatation); and secondly, a proportion of patients (at least 10%) will require placement of a biliary stent at their first ERCP despite biliary clearance. This necessitates a second ERCP, procedures which were not captured in the data from the previous series. This concept was further discussed in the discussion section (page 73).

These conservative estimates give the below estimated proportional success for each arm used for sample size calculations (as described in Figure 13)

No. procedures required for duct clearance	Immediate dilatation arm	Control
1	0.4	0
2	0.4	0.8
3	0.1	0.1
4	0.1	0.1

### Complications

Our estimates used for sample-size calculation are based on complication rates from Ho's previously published series from our institution. These can be seen below (adapted from table 2 in Ho's paper).

<b>Complication</b>	<b>Immediate dilatation (32 pts)</b>	<b>Delayed sphincteroplasty (104 pts)</b>
Bleeding	3 (10%)	1 (1%)
Perforation	0	1 (1%)
Cholangitis	1 (3%)	3 (3%)
Pancreatitis	0	1 (1%)
<b>Total</b>	<b>12.5%</b>	<b>6%</b>

## Appendix E: Full Complication Data Table

Participant	Arm	Complication	ERCP during which the complication occurred	Grade (Cotton criteria)	Comprehensive Complication Index
2	Immediate dilatation	Cholangitis	1	Mild	20.9
3	Immediate dilatation	? aspiration	1	Mild	8.7
10	Control	Pancreatitis	1	Mild	8.7
12	Control	Perforation	2	Severe	42.4
18	Control	Pancreatitis	1	Moderate	20.9
23	Immediate dilatation	Pancreatitis	1	Mild	8.7
31	Immediate dilatation	Bleeding	2	Moderate	26.2
39	Control	Pancreatitis	1	Moderate	8.7
42	Immediate dilatation	Cholangitis	2	Moderate	8.7

Grade calculated as per Cotton criteria (30)

Comprehensive Complication index calculated via online calculator (37)

## Bibliography

1. Conlon K. The gall bladder and bile ducts. In: Norman Williams CB, P Ronan O'Connell, editor. Bailey & Love's Short Practice of Surgery. 26 ed: CRC Press; 2013.
2. Andriulli A, Loperfido S, Napolitano G, Niro G, Valvano MR, Spirito F, et al. Incidence rates of post-ERCP complications: a systematic survey of prospective studies. *American Journal of Gastroenterology*. 2007;102(8):1781-8.
3. AGSE Standards of Practice Committee. Complications of ERCP. *Gastrointestinal Endoscopy*. 2012;75(3):467-73.
4. Classen M, Demling L. [Endoscopic sphincterotomy of the papilla of Vater and extraction of stones from the choledochal duct (author's translation)]. *Dtsch Med Wochenschr*. 1974;99(11):496-7.
5. Kawai K, Akasaka Y, Murakami K, Tada M, Koli Y. Endoscopic sphincterotomy of the ampulla of Vater. *Gastrointestinal Endoscopy*. 1974;20(4):148-51.
6. Enochsson L, Swahn F, Arnelo U, Nilsson M, Lohr M, Persson G. Nationwide, population-based data from 11,074 ERCP procedures from the Swedish Registry for Gallstone Surgery and ERCP. *Gastrointestinal Endoscopy*. 2010;72(6):1175-84, 84 e1-3.
7. Staritz M, Ewe K, Meyer zum Buschenfelde KH. Endoscopic papillary dilatation, a possible alternative to endoscopic papillotomy. *Lancet*. 1982;1(8284):1306-7.
8. Ersoz G, Tekes O, Ozutemiz AO, Gunsar F. Biliary sphincterotomy plus dilation with a large balloon for bile duct stones that are difficult to extract. *Gastrointestinal Endoscopy*. 2003;57(2):156-9.
9. Disario JA, Freeman ML, Bjorkman DJ, Macmathuna P, Petersen BT, Jaffe PE, et al. Endoscopic balloon dilation compared with sphincterotomy for extraction of bile duct stones. *Gastroenterology*. 2004;127(5):1291-9.
10. Weinberg B, Shindy W, Lo S. Endoscopic balloon sphincter dilation (sphincteroplasty) versus sphincterotomy for common bile duct stones (Review). *Cochrane Database of Systematic Reviews*. 2009(1):1-30.
11. Attam R, Freeman M. Endoscopic papillary large balloon dilation for large common bile duct stones. *Journal of Hepatobiliary Surgery*. 2009;16:618 – 23.
12. Katanuma A, Maguchi H, Osanai M, Takahashi K. Endoscopic treatment of difficult common bile duct stones. *Digestive Endoscopy*. 2010;22 Suppl 1:S90-7.
13. Jin PP, Cheng JF, Liu D, Mei M, Xu ZQ, Sun LM. Endoscopic papillary large balloon dilation vs endoscopic sphincterotomy for retrieval of common bile duct stones: A meta-analysis. *World Journal of Gastroenterology*. 2014;20(18):5548-56.

14. Attasaranya S, Cheon YK, Vittal H, Howell DA, Wakelin DE, Cunningham JT, et al. Large-diameter biliary orifice balloon dilation to aid in endoscopic bile duct stone removal: a multicenter series. *Gastrointestinal Endoscopy*. 2008;67(7):1046-52.
15. Meine GC, Baron TH. Endoscopic papillary large-balloon dilation combined with endoscopic biliary sphincterotomy for the removal of bile duct stones (with video). *Gastrointestinal Endoscopy*. 2011;74(5):1119-26.
16. Heo JH, Kang DH, Jung HJ, Kwon DS, An JK, Kim BS, et al. Endoscopic sphincterotomy plus large-balloon dilation versus endoscopic sphincterotomy for removal of bile-duct stones. *Gastrointestinal Endoscopy*. 2007;66(4):720-6; quiz 68, 71.
17. Qian JB, Xu LH, Chen TM, Gu LG, Yang YY, Lu HS. Small endoscopic sphincterotomy plus large-balloon dilation for removal of large common bile duct stones during ERCP. *Pakistan Journal of Medical Science*. 2013;29(4):907-12.
18. Teoh AY, Cheung FK, Hu B, Pan YM, Lai LH, Chiu PW, et al. Randomized trial of endoscopic sphincterotomy with balloon dilation versus endoscopic sphincterotomy alone for removal of bile duct stones. *Gastroenterology*. 2013;144(2):341-5 e1.
19. Kim HG, Cheon YK, Cho YD, Moon JH, Park DH, Lee TH, et al. Small sphincterotomy combined with endoscopic papillary large balloon dilation versus sphincterotomy. *World J Gastroenterol*. 2009;15(34):4298-304.
20. Feng Y, Zhu H, Chen X, Xu S, Cheng W, Ni J, et al. Comparison of endoscopic papillary large balloon dilation and endoscopic sphincterotomy for retrieval of choledocholithiasis: a meta-analysis of randomized controlled trials. *Journal of Gastroenterology*. 2012;47:655-63.
21. Stefanidis G, Viazis N, Pleskow D, Manolakopoulos S, Theocharis L, Christodoulou C, et al. Large balloon dilation vs. mechanical lithotripsy for the management of large bile duct stones: a prospective randomized study. *American Journal of Gastroenterology*. 2011;106(2):278-85.
22. Itoi T, Itokawa F, Sofuni A, Kurihara T, Tsuchiya T, Ishii K, et al. Endoscopic sphincterotomy combined with large balloon dilation can reduce the procedure time and fluoroscopy time for removal of large bile duct stones. *The American Journal of Gastroenterology*. 2009;104:560-5.
23. Kim TH, Oh HJ, Lee JY, Sohn YW. Can a small endoscopic sphincterotomy plus a large-balloon dilation reduce the use of mechanical lithotripsy in patients with large bile duct stones? *Surgical Endoscopy*. 2011;25(10):3330-7.
24. Lin CK, Lai KH, Chan HH, Tsai WL, Wang EM, Wei MC, et al. Endoscopic balloon dilatation is a safe method in the management of common bile duct stones. *Digestive and Liver Disease*. 2004;36(1):68-72.
25. Oh MJ, Kim TN. Prospective comparative study of endoscopic papillary large balloon dilation and endoscopic sphincterotomy for removal of large bile duct stones in patients above 45 years of age. *Scandinavian Journal of Gastroenterology*. 2012;47(8-9):1071-7.

26. Vandervoort J, Soetikno RM, Tham TC, Wong RC, Ferrari AP, Jr., Montes H, et al. Risk factors for complications after performance of ERCP. *Gastrointestinal Endoscopy*. 2002;56(5):652-6.
27. Horiuchi A, Nakayama Y, Kajiya M, Kato N, Kamijima T, Graham DY, et al. Biliary stenting in the management of large or multiple common bile duct stones. *Gastrointestinal Endoscopy*. 2010;71(7):1200-3 e2.
28. Aujla UI, Lee HL, Ryan-Fisher C, Noorullah O, Dwyer LK, Hood SV, et al. Endoscopic papillary large volume balloon dilatation is a safe and effective technique for biliary stone removal independent of timing or size of sphincterotomy. *Gastrointestinal Endoscopy*. 2015;81(5S):AB360-AB1.
29. Ho S, Rayzan D, Fox A, Kalogeropoulos G, Mackay S, Hassen S, et al. Endoscopic sphincterotomy with sphincteroplasty for the management of choledocholithiasis: a single-centre experience. *ANZ Journal of Surgery*. 2017;87(9):695-9.
30. Cotton PB, Lehman G, Vennes J, Geenen JE, Russell RCG, Meyers WC, et al. Endoscopic sphincterotomy complications and their management: an attempt at consensus. *Gastrointestinal Endoscopy*. 1991;37(3):383-93.
31. Kang X, Zheng L, Zeng W, Yang S, Sun H, Zhang R, et al. Risk factors for post-ERCP pancreatitis in high-risk patients receiving post-procedure rectal indomethacin. *Journal of Gastrointestinal Surgery*. 2018.
32. Slankamenac K, Graf R, Barkun J, Puhan MA, Clavien PA. The comprehensive complication index: a novel continuous scale to measure surgical morbidity. *Annals of Surgery*. 2013;258(1):1-7.
33. Slankamenac K, Nederlof N, Pessaux P, de Jonge J, Wijnhoven BP, Breitenstein S, et al. The comprehensive complication index: a novel and more sensitive endpoint for assessing outcome and reducing sample size in randomized controlled trials. *Annals of Surgery*. 2014;260(5):757-62; discussion 62-3.
34. Faul F, Erdfelder E, Lang AG, Buchner A. G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behaviour Research Methods*. 2007;39(2):175-91.
35. Freeman ML. Complications of endoscopic biliary sphincterotomy: a review. *Endoscopy*. 1997;29(4):288-97.
36. Inamdar S, Han D, Passi M, Sejjal DV, Trindade AJ. Rectal indomethacin is protective against post-ERCP pancreatitis in high-risk patients but not average-risk patients: a systematic review and meta-analysis. *Gastrointestinal Endoscopy*. 2017;85(1):67-75.
37. Access Surgery. CCI®-Calculator 2018 [cited 2018 29/7/18]. Available from: [https://www.assessurgery.com/about\\_cci-calculator/](https://www.assessurgery.com/about_cci-calculator/).
38. Owens W, Felts J, Spitznagel E. ASA Physical Status classifications: a study of consistency of ratings. *Anesthesiology*. 1978;49:239-43.
39. Kapral C, Duller C, Wewalka F, Kerstan E, Vogel W, Schreiber F. Case volume and outcome of endoscopic retrograde cholangiopancreatography: results of a nationwide Austrian benchmarking project. *Endoscopy*. 2008;40(8):625-30.

40. Kim HJ, Choi HS, Park JH, Park DI, Cho YK, Sohn CI, et al. Factors influencing the technical difficulty of endoscopic clearance of bile duct stones. *Gastrointestinal Endoscopy*. 2007;66(6):1154-60.
41. Sundaralingam P, Masson P, Bourke MJ. Early precut sphincterotomy does not increase risk during endoscopic retrograde cholangiopancreatography in patients with difficult biliary access: a meta-analysis of randomized controlled trials. *Clinical Gastroenterology and Hepatology*. 2015;13(10):1722-9 e2.
42. Borchert DH, Federlein M, Muller VA, Wagenpfeil S, Eisele RM. Comprehensive complication index for NOTES procedures: results from a randomized controlled trial and comparison to published NOTES complication data. *Surgical Endoscopy*. 2015;29(10):2928-33.
43. Australian Radiation Protection and Nuclear Safety Agency. Ionising radiation and health Yallambie, Victoria, Australia: ARPANSA; 2015 [cited 2018 15/7/18]. Available from: <https://www.arpansa.gov.au/understanding-radiation/radiation-sources/more-radiation-sources/ionising-radiation-and-health>.
44. Pierce RA, Jonnalagadda S, Spitler JA, Tessier DJ, Liaw JM, Lall SC, et al. Incidence of residual choledocholithiasis detected by intraoperative cholangiography at the time of laparoscopic cholecystectomy in patients having undergone preoperative ERCP. *Surgical Endoscopy*. 2008;22(11):2365-72.
45. O'Connor U, Gallagher A, Malone L, O'Reilly G. Occupational radiation dose to eyes from endoscopic retrograde cholangiopancreatography procedures in light of the revised eye lens dose limit from the International Commission on Radiological Protection. *British Journal of Radiology*. 2013;86(1022):20120289.
46. Cotton PB, Garrow DA, Gallagher J, Romagnuolo J. Risk factors for complications after ERCP: a multivariate analysis of 11,497 procedures over 12 years. *Gastrointestinal Endoscopy*. 2009;70(1):80-8.
47. He X, Zheng W, Ding Y, Tang X, Si J, Sun LM. Rectal Indomethacin Is Protective against Pancreatitis after Endoscopic Retrograde Cholangiopancreatography: Systematic Review and Meta-Analysis. *Gastroenterol Res Pract*. 2018;2018:9784841.
48. Luo H, Zhao L, Leung J, Zhang R, Liu Z, Wang X, et al. Routine pre-procedural rectal indometacin versus selective post-procedural rectal indometacin to prevent pancreatitis in patients undergoing endoscopic retrograde cholangiopancreatography: a multicentre, single-blinded, randomised controlled trial. *Lancet*. 2016;387(10035):2293-301.
49. Kim JH, Yang MJ, Hwang JC, Yoo BM. Do balloons larger than 15 mm in size increase the risk of adverse events in endoscopic large balloon dilation? *Gastrointestinal Endoscopy*. 2015;81(5S):AB360-1.
50. Poh BR, Cashin PA, Croagh DG. Impact of jaundice on outcomes following emergency laparoscopic cholecystectomy in patients with choledocholithiasis. *World Journal of Surgery*. 2018;42(10):3158-64.
51. Board of Radiation Effects Research Division on Earth and Life Sciences National Research Council of the National Academies. Health risks from exposure to low levels of ionizing radiation: BEIR VII Phase 2. Washington, DC: National Academies Press; 2006.



52. Hanley M, Koonce J, Bradshaw M. Calculate your risk 2018 [cited 2018 17/7/18]. Available from: <http://www.xrayrisk.com/calculator/calculator.php>.
53. Cohen RV, Aldred MA, Paes WS, Fausto AM, Nucci JR, Yoshimura EM, et al. How safe is ERCP to the endoscopist? *Surgical Endoscopy*. 1997;11(6):615-7.
54. Heyd RL, Kopecky KK, Sherman S, Lehman GA, Stockberger SM. Radiation exposure to patients and personnel during interventional ERCP at a teaching institution. *Gastrointestinal Endoscopy*. 1996;44(3):287-92.
55. Kim E, McLoughlin M, Lam EC, Amar J, Byrne M, Telford J, et al. Retrospective analysis of radiation exposure during endoscopic retrograde cholangiopancreatography: critical determinants. *Canadian Journal of Gastroenterology*. 2011;25(10):555-9.
56. Kachaamy T, Harrison E, Pannala R, Pavlicek W, Crowell MD, Faigel DO. Measures of patient radiation exposure during endoscopic retrograde cholangiography: beyond fluoroscopy time. *World Journal of Gastroenterology*. 2015;21(6):1900-6.
57. Uradomo LT, Lustberg ME, Darwin PE. Effect of physician training on fluoroscopy time during ERCP. *Digestive Diseases and Sciences*. 2006;51(5):909-14.
58. Enochsson L, Blohm M, Sandblom G, Jonas E, Hallerback B, Lundell L, et al. Inversed relationship between completeness of follow-up and coverage of postoperative complications in gallstone surgery and ERCP: a potential source of bias in patient registers. *BMJ Open*. 2018;8(1):e019551.
59. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Annals of Surgery*. 2004;240(2):205-13.