OCCASIONAL VIEWPOINT

Evaluating ERCP is important but difficult

P B Cotton

Gut 2002;51:287–289

ERCP is a valuable technique now practised widely throughout the world. It revolutionised the diagnosis and management of benign and malignant biliary and pancreatic diseases in the 1970s and 1980s. However, recent developments have highlighted the need for detailed evaluation of current ERCP practice. This review is based on a presentation to a recent NIH “state of the science” conference on ERCP, and refers to an article which appears in this issue of *Gut* from researchers in Hong Kong who report on a randomised controlled trial of endoscopic sphincterotomy in acute cholangitis.

Endoscopic retrograde cholangiopancreatography (ERCP) is a valuable technique now practised widely throughout the world. Without question it revolutionised the diagnosis and management of benign and malignant biliary problems in the late 1970s and 1980s. It was accepted rapidly, with very few studies, because it was so obviously preferable to surgery which then carried substantial risks. The situation now is very different. Imaging methods have proliferated in number and sophistication, so that we hear frequently that diagnostic ERCP is obsolete, or at least obsolescent. Equally important, surgical techniques have improved enormously, with increased emphasis on minimally invasive techniques wherever possible, and greatly helped by improved anaesthesia, perioperative care, and intensive care. Operative mortality values have dropped so low (partly because many of the higher risk patients are triaged appropriately to endoscopic intervention) that mortality is no longer the main comparative driver in discussions of the relative merits of different approaches. Comparisons become more difficult as the possible outcome parameters proliferate (for example, costs, quality of life measures, patient preferences) and become somewhat softer. But we do need good data, especially as endoscopists attempt to extend their roles into more speculative clinical contexts, such as sphincter of Oddi dysfunction (SOD) and pancreatitis.

These considerations led the American Society for Gastrointestinal Endoscopy (ASGE) to request, and the National Institutes of Health (NIH) recently to hold, a consensus conference entitled “State of the science conference on ERCP”. Like other NIH consensus conferences, this relied on a wise panel of physicians (and a few lay representatives) who were not personally involved in the topic. They heard presentations from 15 experts in ERCP, radiology, and surgery. They also commissioned a special literature review which used strict quality criteria, focusing especially on

controlled trials. It became brutally obvious that there have been very few relevant randomised studies (at least of any quality). The panel made many useful observations and conclusions.¹ Not surprisingly, one main recommendation was that more randomised studies were needed.

“Why is our ERCP practice not based firmly on the results of randomised controlled trials?”

We are forced to ask the painful question—why is our ERCP practice not based firmly on the results of randomised controlled trials? Are most endoscopists (and surgeons) simply incapable of doing good science? There are certainly plenty of examples of flawed studies, including many of my own. But it is rather more complicated than that.

In this issue of *Gut*,² researchers from the Queen Mary Hospital (Hong Kong) report a randomised controlled trial of endoscopic sphincterotomy in acute cholangitis [see page 245]. The study asked a simple question—that is, whether or not to do a sphincterotomy when you do not find any stones in the duct during ERCP in a patient with cholangitis (and gall bladder stones). After randomising 111 patients, the sphincterotomy group appeared to have shorter durations of fever and hospital admission but also more cholangitis in follow up. It is difficult to evaluate the results as the authors do not define the primary hypothesis nor provide a power calculation. They concluded that the apparent slight benefits might not justify the increased risk of sphincterotomy, although the complication rates were very low. Thus the study really did not answer the question. Sadly, even a definite answer either way would be of little relevance to Western endoscopists who have different patients and practices. It is particularly striking that none of the 111 patients with cholangitis and gall bladder stones were referred for cholecystectomy. This study is an example of many that have been performed to answer questions that are essentially technical, involving a choice between two actions during a single procedure. The sister group at the Prince of Wales Hospital across Hong Kong harbour has done many analogous studies comparing different haemostatic methods during endoscopy in acute bleeding. Similar trials have compared two different current modalities during sphincterotomy, and the choice between banding

Correspondence to:
P B Cotton, Digestive
Disease Center, Medical
University of South
Carolina, PO Box
250327, Charleston, SC
29425, USA;
cottonp@usc.edu

Accepted for publication
19 February 2002

Abbreviations: ERCP, endoscopic retrograde cholangiopancreatography; SOD, sphincter of Oddi dysfunction; NIH, National Institutes of Health; ESWL, extracorporeal shock wave lithotripsy.

and sclerotherapy in variceal bleeding. These studies are interesting, and are relatively easy to perform (as they are only minor variants during a standard procedure), but the really crucial questions are on a different plane altogether.

"Do certain ERCP treatments have any value at all?"

What patients and NIH panels really want to know are, firstly, do certain ERCP treatments have any value at all?—do they work? And, secondly, are they "better" than available methods (especially surgery)? The first question could be addressed by a placebo controlled study, and the second by a head to head comparison with surgery. Some might argue for trying to answer both questions at once with randomisation to placebo, endoscopy, and surgery (as surgery cannot claim to be a gold standard). While not questioning the desirability of having the data that such studies could generate, they are extremely difficult to mount and few will ever happen. I have argued forcibly elsewhere that our baseline knowledge can be advanced considerably using simpler cohort studies, as several groups have illustrated.³⁻⁵

Those determined to pursue randomised trials of major interventions will encounter many barriers. Firstly, it is often difficult to define a tight cohort of participants, especially in the contexts which most need evaluation (that is, SOD and pancreatitis). Patients come in all shapes and sizes, with different suspicions and stages of disease, levels of disability, and perspectives on appropriate outcomes. While exact entry criteria can be defined, it is obvious that the results apply only to that small group. Secondly, ERCP (like surgery) is not a "pure" intervention. Results certainly depend on the particular endoscopist and the team involved. Thirdly, technologies continue to evolve. It is a pity to spend many years and dollars coming to a conclusion that is obsolete before publication. Fourthly, the risks of ERCP (especially in the context of suspected SOD) make it difficult to justify a placebo or sham treatment. Fifthly, potential participants may be reluctant to consider randomisation between interventions which differ so much in their nature, and perceived burden. Why risk needing two weeks in hospital (or worse) when a day case procedure may do the trick? Why don't we just try the ERCP treatment first, Doctor? That is a reasonable question provided there is some evidence that the endoscopic treatment is indeed sometimes successful. There are also important practical issues related to the numbers of patients needed for trials, referral patterns, and patient expectations. Centres with the most patients with suspected SOD and pancreatitis gain that experience because referring primary physicians and gastroenterologists think that the experts know what they are doing. Patients arrive expecting to receive the special treatment that their local physician has recommended so glowingly (and which our web sites make sound so attractive). It is surely mainly our fault if others believe that we have the answer but it is also difficult for experts to admit ignorance. Many interventionists have egos which need to be fed regularly. That is not an adequate reason for knee jerk application of the expected treatment (and failing to enter patients into important studies), but the pressures are difficult to resist, at least while maintaining the confidence of our patients and our referral sources. The problem should be addressed by careful education of referrers to ensure that patients arrive expecting to be studied and managed thoughtfully and appropriately, and not automatically.

"The risks of ERCP make it difficult to justify a placebo or sham treatment"

The clash between our duty to individual patients and to science has been well discussed.^{6,7} Randomisation is appropri-

ate only in the state of equipoise—that is, when the planning protagonists honestly do not know which procedure is preferable. It may be difficult to get endoscopists and surgeons to agree on this point. When they do, the tested cohort may be only a small part of the overall spectrum of the disease and the results will have no generalisability. Ten years ago the same group at Queen Mary Hospital in Hong Kong reported a seminal randomised trial showing that ERCP was much safer than surgery in patients with acute cholangitis.⁸ This study worried me at the time because the power calculation was based on the number of likely excess deaths in the surgical arm⁹—a prediction that was adequately fulfilled.

If all of these problems are overcome, there remains the issue of financial support. Multicentre long term trials are very expensive, especially if one of the interventions is judged to be experimental (or a sham treatment). For instance, a study of endoscopic treatment in calcific chronic pancreatitis must include extracorporeal shock wave lithotripsy (ESWL), maybe several times. How can this be done in the USA if ESWL is neither FDA approved nor reimbursed?

Some who have read this far may question my commitment to scientific evaluation. Documenting the difficulties does not mean that I am disinterested—but I have learned from some experience. From the Middlesex Hospital in London we published two randomised studies addressing palliation of malignant obstructive jaundice. We performed the only study comparing endoscopic with percutaneous stenting,¹⁰ and the largest randomised study comparing ERCP stenting and surgical bypass.¹¹ Both studies burnished some resumes at the time (and generated discussion at the NIH conference) but the results are quite irrelevant now, and actually had little impact at the time. The percutaneous/endoscopic comparison was outdated rapidly as radiological techniques developed. The surgery/stenting study was more interesting. The immediate relief of jaundice was similar in the two groups but the risks in the stented group were (not surprisingly) substantially less. However, as follow up continued, the stented group also experienced significantly more recurrent jaundice due to stent occlusion. These results were easy to predict beforehand. The study was criticised for the high complication rates of surgery, with the suggestion that surgical expertise was suboptimal. Since the study was performed in the mid 1980s, this is a reasonable criticism, but it is also true that stenting has improved substantially since that time. For example, expandable metal stents have changed the metrics of the debate.

Another problem is that such randomised trials are designed to see which of two treatments is "better", but this has to be put in context. There is no "one size fits all". There is a wide spectrum of patients with malignant obstructive jaundice with varying degrees of operability and resectability (fig 1). Surgery is clearly appropriate for healthy (operable) patients believed to have lesions which are resectable (fig 1, box 3), or maybe resectable (box 6), and for patients with resectable lesions who are probably fit for surgery (box 2). Equally clearly, endoscopic stenting is preferable in patients who are not operative candidates (boxes 1, 4, 7), and in most who are not resectable and marginally operable (box 8). The position of equipoise (and debate) concerns only the relatively few patients who are clearly both operable and unresectable (box 9), and those who may be resectable and may be operable (box 5). A new randomised trial within that small cohort would be of limited general value, and there would be a risk that the results would be extrapolated inappropriately to other patients. Also, what would be the primary outcome parameter—the number of procedures, total hospital days, quality of life, costs, survival, or some magic combination representing our idea of "better"? It might be of greater practical interest to study patient's attitudes to these different outcomes. In the context of malignant jaundice, which patients would choose the simpler, easier, cheaper, and safer (but less effective) stenting approach, and who would opt for

Operable?	No	1 Stent	4 Stent	7 Stent
	Maybe	2 Surgery	5 ?	8 Stent
	Yes	3 Surgery	6 Surgery	9 ?
		Yes	Maybe	No
		Resectable?		

Figure 1 Varieties of patients with malignant obstructive jaundice, and treatment options.

one time surgery, probably more effective, but certainly more risky and initially painful? What drives those preferences? Which brings us back to the real question that is asked by each individual patient—which treatment is best for me, Doctor—with my disease stage, symptoms, prognosis, and attitudes—and probable chosen therapist?

There is another important outcome dilemma in studies which do not involve malignancy—for example, management of chronic pancreatitis. What is the time frame of evaluation? Endoscopic (and surgical) treatments are rarely curative in this context. Is the outcome measured at one year or five years, or even longer?

Cohort studies can be designed to clarify the predictors of good and bad outcomes by all possible alternative techniques. However, they do require much greater discipline and objectivity than most currently reported. Firstly, different interventionists (that is, endoscopists and surgeons) must use the same language in defining and stratifying patients.¹² This is particularly relevant in terms of disease stage, activity, expertise, and operative risk. Secondly, the appropriate outcome measures must be agreed—not only by the protagonists but also by patients. Another problem is objectivity. Most cohort studies of endoscopic (and surgical) procedures have been evaluated by the protagonists of these techniques, which seems naive at best. We cannot be both prosecutor and judge. Surely we need to involve objective scientific “skeptics” to ensure that the study data are believable. Thus a cohort study of endoscopic intervention might include a surgeon in the planning and execution, and vice versa.

Although we can hope to see better studies in the future from interested individuals and groups, it is my contention

My recommendations for progress in this area are:

- Get pancreatobiliary interventionists (surgeons, endoscopists, and interventional radiologists) to agree on a common language for describing patients, interventions, and outcomes.
- Use these definitions in careful prospective cohort studies using unbiased observers.
- Use the resulting data to define which randomised trials are really necessary and worthwhile.
- Challenge national organisations to provide the funding.

that real progress requires a concerted approach at a national level. The appropriate professional organisations must embrace and support this agenda. Such an approach has been developed and is productive in evaluating new cancer treatments.

REFERENCES

- 1 www.nih.gov. Draft statement ERCP for diagnosis and therapy January 14–16, 2002.
- 2 Hui C-K, Lai K-C, Wong W-M, *et al*. A randomised controlled trial of endoscopic sphincterotomy in acute cholangitis without common bile duct stones. *Gut* 2002;**51**:245–7.
- 3 Freeman ML, DiSario JA, Nelson DB, *et al*. Risk factors for post-ERCP pancreatitis: a prospective, multicenter study. *Gastrointest Endosc* 2001;**54**:425–34.
- 4 Freeman ML, Nelson BD, Sherman S, *et al*. Complications of endoscopic biliary sphincterotomy. *N Engl J Med* 1996;**335**:909–18.
- 5 Davis WZ, Cotton PB, Arias R, *et al*. ERCP and sphincterotomy in the context of laparoscopic cholecystectomy: academic and community practice patterns and results. *Am J Gastroenterol* 1997;**92**:597–601.
- 6 McDermott WM, Amoroso CS. Sounding board of mice but not men. *N Engl J Med* 1991;**324**:1585–660.
- 7 Passamani E. Clinical trials—are they ethical? *N Engl J Med* 1991;**324**:1589–92.
- 8 Lai ECS, Mok FPT, Tan ESY, *et al*. Endoscopic biliary drainage for severe acute cholangitis. *N Engl J Med* 1992;**326**:1582–6.
- 9 Cotton PB. Therapeutic gastrointestinal endoscopy. Problems in proving efficacy. *N Engl J Med* 1992;**326**:1626–8.
- 10 Speer AG, Cotton PB, Russell RCG, *et al*. Randomized trial of endoscopic versus percutaneous stent insertion in malignant obstructive jaundice. *Lancet* 1987;**2**:57–62.
- 11 Smith AC, Dowsett JF, Russell RCG, *et al*. Randomised trial of endoscopic stenting versus surgical bypass in malignant low bile duct obstruction. *Lancet* 1994;**344**:1655–60.
- 12 Cotton PB. Randomization is not the (only) answer: a plea for structured objective evaluation of endoscopic therapy. *Endoscopy* 2000;**32**:402–5.