Evaluating ERCP is important but difficult

P B Cotton

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 Imaging methods have proliferated in number and sophistication, so that we hear frequently that diagnostic ERCP is obsolete, or at least obsolete. 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. 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, 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...
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. 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 expertise 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. Randomisation is appropri-
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