(hematuria in a small percent, postprocedure colic, increase ALT or AST in 8% of patients), these quickly revert to normal. It is precisely because any pain or discomfort associated with lithotripsy resolves so quickly and the patients immediately resume their previous activities that the procedure is so attractive to them.

ESWL will not be a viable alternative to cholecystectomy unless it is highly effective, ie, the gallbladder becomes stone-free, and unless stone recurrence is infrequent in successfully treated patients. Efficacy with ESWL is improved by appreciation that the optimal fragmentation will be achieved in patients with one to three stones less than 20 mm in size and also that an institution undergoes a "learning curve" with the procedure, with more successful treatments being achieved as experience is gained. The rate and rapidity of the patients becoming stone-free has increased at all of the study centers as we have gained experience.

Recurrence of cholesterol stones in treated patients appears to be confined to a subset, generally occurs early (one to three years), and may be prevented by treatment with low-dose intermittent bile acid therapy or aspirin (1, 2). We certainly agree with Dr. Flick that alternative therapies will not be attractive if recurrent treatments in many patients are required. Efforts are underway to better define the patient population who are not at risk for recurrence and to develop methods to prevent recurrence in those who are.

Dr. Flick raises logical concerns about the cost effectiveness of the procedure and potential inappropriate use. Unfortunately, we too often equate charges for the procedure with its true cost. Reimbursement to hospitals is increasingly disconnected from charges, and the trend is likely to continue. We need to determine the true cost of different approaches since it is on this basis that economic decisions will be made in the future.

Finally, we need to examine not just new, but already available technology and treatments relative to their appropriate application. It seems clear that a significant fraction of the 500,000 cholecystectomies performed in the United States each year are done without appropriate indications. Since cholecystectomy, like lithotripsy, is performed in patients for abdominal pain, and since recent studies have demonstrated that in approximately 50% of patients the precholecystectomy pain continues after the operation (3), outcome and effectiveness studies should be applied to both procedures.

Dr. Flick raises many excellent points that have commonly been discussed in regard to gallstone lithotripsy. However, we believe that there is a group of patients for whom lithotripsy, followed by oral dissolution treatment with ursodeoxycholic acid, is both safe and effective therapy and other high-risk patients for whom it offers a definite safety advantage. To deny the technology to those for whom it is an appropriate modality by using standards for this procedure not applied to the alternatives is inconsistent and unfair to those patients.

DAVID A. BURNETT, MD
Department of Medicine
University of Nebraska Medical Center
Omaha, Nebraska

NORTON J. GREENBERGER, MD
Department of Medicine
University of Kansas Medical Center
Kansas City, Kansas

REFERENCES

EXOGENOUS PROSTAGLANDINS, ALKALINE SECRETION, AND PROTECTION OF DUODENAL MUCOSA

To the Editor:

In the November 1989 issue of Digestive Diseases and Sciences, Leung et al (1) presented some interesting data concerning the mechanism of the protective effect of prostaglandins on the duodenal mucosa. It has been shown before that exogenous prostaglandins stimulate duodenal mucosal bicarbonate secretion (2), and others have shown that prostaglandins can protect gastrointestinal mucosa against several types of damaging agents including HCl, but Leung et al are the first to show that prostaglandins stimulate bicarbonate secretion and at the same time protect the mucosa. Apart from the title of the paper of Leung et al being somewhat misleading (it states that a causality between the two phenomena mentioned above has been shown
while in fact merely coexistence between them has been shown), the paper contains some serious statistical errors. To evaluate the protective effect of prostaglandins on the duodenal mucosa the authors have constructed a grading scale for the histological damage on the villi of the mucosa (grade 0–IV). This scale is a rank scale. Grade IV represents a more serious damage than grade III, which again represents a more serious damage than grade II, and so on. The distance between one grade and the next grade is, however, not the same throughout the scale (it is not a ratio/interval scale), and therefore it is incorrect to calculate the mean lesion grade for a group of villi, the standard deviation and the standard error of the mean, as the authors have done. Furthermore the authors use an unpaired $t$ test to test for difference in acid-induced lesions between untreated rats and rats treated with prostaglandins, even though this test requires normally distributed data measured by a ratio/interval scale. The error made by using a parametric test ($t$ test) on nonnormally distributed data is usually small when the number of data is large. As only six rats were examined in each group and as the difference between the lesions of the two groups was small, the risk of making a false conclusion by using a $t$ test on data that are neither normally distributed nor measured by a ratio/interval scale is considerable. As it is crucial to the conclusions of the paper that a difference in lesions between the nontreated rats and the prostaglandin-treated rats can be demonstrated, it is very important that the data are treated correctly. A very easy way to do this is to calculate the median lesion score for each rat and then use the Mann-Whitney U test for unpaired data to test for difference between the median lesion score of the two groups. As the study is very interesting and otherwise very well performed, I hope that the authors will reevaluate their data using correct statistical tests.

MARK A. AINSWORTH, MD
Department of Medical Gastroenterology
Odense University Hospital
DK-5000 Odense C, Denmark

REFERENCES


Response by Terry J. Reedy:

Studies of mucosal damage and protection require a reasonably objective and reproducible measure of mucosal state. For several studies of surface mucosa in the rat corpus, antrum, and duodenum, Dr. Paul Guth and associates such as Dr. Felix Leung have developed and used a measure based on microscopic examination of a histologic cross section for each rat. Fifty units of each slide are scored by an observer unaware of the rat's treatment. The units are eyepiece reticle distance in the stomach and villi in the duodenum. The scores are 0 (normal) to 4 (complete destruction), with definitions of intermediate damage scores 1, 2, and 3 adjusted for each type of mucosa. There are two statistical questions: (1) how do we combine or reduce the 50 unit scores within each rat to produce a composite mucosal damage score for the rat as a whole; and (2) how do we compare the resulting damage scores between rats?

Guth's group has used two types of composite score: percentage of unit scores at or above a threshold level (0–100 in steps of 2) and mean unit score (0–4 in steps of 0.02). A threshold percentage is clearly interpretable, but the best threshold for showing differences is specific to each comparison and any threshold percentage is insensitive to variations in score distribution above and below the threshold value chosen. The mean unit score is more sensitive to distribution differences and is usable for all studies, but the interpretation is less obvious and the summation of ranked scores is open to criticism. (The same is true of most composite measures and even "mean pH"). However, the 0–4 scale is roughly linear-logarithmic and the mean of 50 unit scores can be regarded as a mucosa score giving a composite representation of the typical amount of mucosal damage. The important question is whether the outcome adequately captures this vague, unstandardized concept and is adequately sensitive to changes induced by experimental manipulation. Cross-checks with visual assessments and the outcome of several studies indicate that the mucosal damage score usually meets both criteria.

For comparing treatment groups composed of whole animals, the composite mucosal damage measures, regarded as the outcome of a black-box