1 Introduction

John WD McDonald[1], *Andrew K Burroughs*[2], *Brian G Feagan*[1] *and M Brian Fennerty*[3]

[1]Robarts Clinical Trials, Robarts Research Unit, University of Western Ontario, London, Ontario, Canada
[2]The Royal Free Sheila Sherlock Liver Centre, Royal Free Hospital, *and* University College London, London, UK
[3]Oregon Health and Science University, Division of Gastroenterology and Hepatology, Ontario Oregon, USA

Over the past three decades the emergence of evidence-based medicine (EBM) has had a substantial impact on clinical practice. In the first half of the twentieth century, diagnostic tests or treatments, usually based on a strong scientific rationale and experimental work in animals, were routinely introduced into clinical care without good scientific proof of efficacy in people. Some of these interventions, such as gastric freezing for the treatment of ulcers and penicillamine therapy for primary biliary cirrhosis, were ultimately shown to be ineffective and harmful [1, 2]. There is little doubt that the widespread acceptance by physicians of unproved treatments has been detrimental to the well-being of many patients.

Fortunately, the need for a more critical approach to medical practice was recognized. In 1948 the first randomized controlled trial (RCT) in humans was carried out under the direction of the British Medical Research Council [3]. Epidemiologists and statisticians, notably Sir Richard Doll and Sir Bradford Hill, provided scientific leadership to the medical community, which responded with improvements in the quality of clinical research. The use of randomized allocation to control for confounding variables and to minimize bias was recognized as invaluable for conducting valid studies of treatments. The initiation of these landmark experiments defined a new era in clinical research; the RCT soon became the benchmark for the evaluation of medical and surgical interventions. Gastroenterologists played an important part in these early days. In 1955, Professor Sidney Truelove conducted the first randomized trial in the discipline of gastroenterology [4]. He and his colleagues proved that cortisone was more effective than a placebo for the treatment of ulcerative colitis. As noted in Chapter 12, this treatment has stood the test of time. The ascendancy of the RCT was accompanied by a call for greater scientific rigor in the usual practice of clinical medicine. Strong advocates of the application of epidemiological principles to patient care emerged and found a growing body of support among clinicians.

As the number of randomized trials grew to the point of becoming unmanageable, it was recognized that there was a need to provide summaries of the evidence provided by these trials for the use of practitioners, who frequently lack both time and expertise to consult the primary research. Busy clinicians may consult local experts, with the tacit assumption that they will make recommendations based on evidence. Liberati and colleagues provided evidence that this approach led to inappropriate care for many women with breast cancer [5]. Subsequently, convincing evidence became available through the work of Antman *et al.* and of Mulrow that the conventional review article and the traditional textbook chapter are seldom comprehensive, and are frequently biased [6, 7]. More recently, Jefferson reinforced this conclusion on the basis of a survey concerning recommendations for vaccination for cholera, which appeared in editorials and review articles [8]. He pointed out that authors of editorials and reviews frequently resort to the "desk drawer" technique, pulling out evidence with which they are very familiar, but failing to assemble and review all of the evidence in a systematic way.

In the UK, Archie Cochrane, as early as 1979, made a compelling case that there was a need to prepare and maintain summaries of all randomized trials [9]. Cochrane's challenge to the medical community to use scientific methods to identify, evaluate and systematically summarize the world's medical literature pertaining to all health care interventions is now being met. From its inception in 1993, the electronic database prepared by the volunteer members of the Cochrane Collaboration and published as the *Cochrane Library* has grown exponentially [10]. Systematic reviews and especially Cochrane reviews are now widely used by clinicians in the daily practice of medicine, by researchers and by the public. Accordingly, data from systematic reviews published in the *Cochrane Library*

Evidence-Based Gastroenterology and Hepatology, 3rd edition.
J. McDonald, A.K. Burroughs, B. Feagan, and M.B. Fennerty. © 2010 Blackwell Publishing Ltd

are featured prominently in several chapters in *Evidence-based Gastroenterology and Hepatology*. Unfortunately, coverage in the *Cochrane Library* of topics in gastroenterology and hepatology is still far from complete.

Several other clinical epidemiologists played important roles in the evolution of evidence-based medicine. Beginning in the 1970s, David Sackett encouraged practicing physicians to become familiar with the basic principles of critical appraisal. Criteria developed by Sackett and others for the evaluation of clinical studies assessing therapy, causation, prognosis and other clinical topics were widely published [11, 12]. His text, *Clinical Epidemiology: a Basic Science for Clinical Medicine*, co-authored by colleagues Gordon Guyatt, Brian Haynes and Peter Tugwell, introduced many physicians to the concepts of EBM [13]. In the USA, Alvin Feinstein called attention to the need for increased rigor in the design and interpretation of observational studies and explored the scientific principles of diagnostic testing [14, 15]. Among gastroenterologists, Thomas Chalmers, a strong, early advocate for the RCT [16], was responsible for introducing gastroenterologists and others to the importance of randomized trials in gastroenterology and hepatology and to the concept of systematic reviews and meta-analysis as means of summarizing data from these studies [17, 18].

Despite the opposition of some, the popularity of EBM continues to grow [19]. Although the explanations for this phenomenon are complex, one factor is that many practitioners recognize that ethical patient care should be based on the best possible evidence. For this, and other reasons, the fundamental concept behind EBM – the use of the scientific method in the practice of clinical medicine – has been widely endorsed by medical opinion leaders, patients and governments.

What is evidence-based gastroenterology and hepatology?

Evidence-based gastroenterology and hepatology is the application of the most valid scientific information to the care of patients with gastrointestinal and hepatic diseases. Physicians who treat patients with digestive diseases must provide their patients with the most appropriate diagnostic tests, the most accurate prognosis and the most effective and safe therapy. To meet this high standard individual clinicians must have access to and be able to evaluate scientific evidence. Although many practitioners argue that this has always been the standard of care in clinical medicine, a great deal of evidence exists to the contrary. Wide variations in practice patterns among physicians have been documented for many treatments, despite the presence of good data from widely publicized RCTs and the promotion of practice guidelines by content experts. For example, Scholefield *et al.* carried out a survey of British surgeons who were questioned regarding the performance of screening colonoscopy for colon cancer [20]. Although this study was done in 1998 (after publication of the results of the RCTs described in Chapter 18 which demonstrated a benefit of this practice), many of these physicians failed to make appropriate recommendations for screening patients at risk. What is the explanation for this finding? One possibility is that many clinicians rely for information on their colleagues, on local experts, or on review articles or textbook chapters that are not based on the principles of EBM.

Two important points about EBM should be emphasized. First, use of the principles of EBM in the management of patients is complementary to traditional clinical skills and will never supersede the recognized virtues of careful observation, sound judgment and compassion for the patient. It is noteworthy that many good doctors have intuitively used the basic principles of EBM; hence the promotion of such well-known clinical aphorisms as "go where the money is" and "do the last test first". Knowledge of EBM enables physicians to understand why these basic rules of clinical medicine are valid through the use of a quantitative approach to decision making. This paradigm can in no way be considered detrimental to the doctor-patient relationship.

Second, although RCTs are the most valuable source of data for evaluating health care interventions, other kinds of evidence must frequently be used. In some instances, most obviously in studies of causation, it is neither possible nor ethical to conduct RCTs. Here, data from methodologically rigorous observational studies are extremely valuable. A dramatic example was the demonstration by several authors (quoted in Chapter 27) that the relative risk of hepatocellular carcinoma in chronic carriers of the hepatitis B virus is dramatically higher than in persons who are not infected. Although these data are observational, the strength of the association is such that it is exceedingly unlikely that a cause other than hepatitis B virus is responsible for the development of cancer in these people. Case-control studies are especially useful for studying rare diseases and for the initial development of scientific hypotheses regarding causation. The etiological role of non-steroidal anti-inflammatory drugs in the development of gastric ulcer was recognized using this methodology [21]. Finally, case series can provide compelling evidence for the adoption of a new therapy in the absence of data from RCTs, if the natural history of the disease is both well characterized and severe. An example is the identification of orthotopic liver transplantation as a dramatically effective intervention for patients with advanced liver disease.

Box 1.1 shows a generally agreed approach to ranking the strength of evidence that arises from various types of studies of health care interventions, and this system is used throughout the book. This ranking of evidence has

Box 1.1 Grading of recommendations and levels of' evidence used in *Evidence-based Gastroenterology and Hepatology*

Grade A
Level 1a
- Evidence from large randomized clinical trials (RCTs) or systematic reviews (including meta-analyses) of multiple randomized trials which collectively have at least as much data as one single well-defined trial.

Level 1b
- Evidence from at least one "All or none" high quality cohort study; in which *all* patients died/failed with conventional therapy and some survived/succeeded with the new therapy (e.g. chemotherapy for tuberculosis, meningitis, or defibrillation for ventricular fibrillation): or in which many died/failed with conventional therapy and *none* died/failed with the new therapy (e.g. penicillin for pneumococcal infections).

Level 1c
- Evidence from at least one moderate sized RCT or a meta-analysis of small trials which collectively only has a moderate number of patients.

Level 1d
- Evidence from at least one RCT.

Grade B
Level 2
- Evidence from at least one high quality study of non-randomized cohorts who did and did not receive the new therapy.

Level 3
- Evidence from at least one high quality case control study.

Level 4
- Evidence from at least one high quality case series.

Grade C
Level 5
- Opinions from experts without reference or access to any of the foregoing (for example, argument from physiology, bench research car first principles)

appeared in a number of publications; we have chosen to reproduce it from *Evidence-based Cardiology*, along with the system used by its editors, Yusuf *et al.*, for making recommendations on the basis of these levels of evidence [22]. As mentioned in Box 1.1, throughout this book recommendation grades appear as **A** or **A1a**.

Clinical decision making in gastroenterology and hepatology

Clinical decision making by gastroenterologists usually falls into one of the following categories:

- Deciding whether to apply a specific diagnostic test in arriving at an explanation of a patient's problem, or determining the status of the patient's disease.
- Offering a prognosis to a patient.
- Deciding among a number of interventions available for managing a patient's problem. In this category, the first question is "Does a given intervention do more good than harm?" The second is "Does it do more good than other effective interventions?" The third is "Is it more or less cost-effective than other interventions?"

A comprehensive approach would incorporate many different types of evidence (e.g. RCTs, non-RCTs, epidemiologic studies and experimental data), and examine the architecture of the information for consistency, coherence and clarity. Occasionally, the evidence does not completely fit into neat compartments. For example, there is strong (A1a) evidence through very large randomized trials that fecal occult blood testing on an annual or semi-annual basis modestly reduces mortality from colon cancer in a population at average risk for this disease. The evidence that direct examination of the colon at intervals of five to ten years results in even greater benefit has been derived only from case control studies (B3). Physicians, patients and policy advisers should have both levels of evidence available to make informed decisions.

Recommendation grades appear either within the text, for example **A** and **A1a** or within a table in the chapter.

The grading system clearly is only applicable to preventative or therapeutic interventions. It is not applicable to many other types of data such as descriptive, genetic or pathophysiologic.

Application of a diagnostic test

Example: A four-year-old child is experiencing diarrhea and has a positive family history of celiac disease. Should a serological test for antiendomysial antibody (EMA) be done?

Chapter 10 includes an extensive treatment of this topic with a summary of studies (see Table 10.1) that included various groups of patients with a greater or lesser probability of having celiac disease (ranging from patients with gastrointestinal symptoms to patients in whom celiac disease was suspected on clinical grounds). Several studies listed in Table 10.1 and the study of Cataldo *et al.* [23] are relevant to this patient.

When evaluating this test the reader may wish to adopt the approach of Kitching *et al.* for deciding on the clinical usefulness of a diagnostic test (Figure 1.1) [24].

The criteria listed in Figure 1.1 for validity of a diagnostic test were clearly met in Cataldo's study. In Chapter 10 Gregor and Say explore the utility of the test and point out that tests with high positive likelihood ratios (LR > 10) and

- **Are the study results valid?**

1 Was there an independent blind comparison (or unbiased comparison) with a reference ("gold") standard of diagnosis?
2 Was the diagnostic test evaluated in an appropriate spectrum of patients (like those seen in the reader's practice)?
3 Was the reference standard applied regardless of the diagnostic test result?

- **What are the results?**

Cataldo F, Ventura A, Lazzari R *et al*. Antiendomysium antibodies and celiac disease: solved and unsolved questions. An Italian multicentre study. *Acta Paediatr* 1995;**84**:1125–31.
A study of IgA endomysium antibodies (EMA) in 1485 children with gastrointestinal disease (688 with celiac disease confirmed by intestinal biopsy)

Results for antiendomysial antibody (EMA) test

	No. of patients with biopsy proven celiac disease		
	Present	Absent	Totals
EMA positive	645	20	665
	a	b	a+b
EMA negative	c	d	c+d
	43	777	810
	a+c	b+d	a+b+c+d
Totals	688	797	1485

Sensitivity = a/(a + c) = 645/688 = 0·94
Specificity = d/(b + d) = 777/797 = 0·97
Likelihood ratio (positive result) = sensitivity/(1–specificity) = 0·94/(1–0·97) = 31
Likelihood ratio (negative result) = (1–sensitivity)/specificity = (1–0·94)/0·97 = 0·06
Positive predictive value = a/(a + b) = 645/665 = 0·97
Negative predictive value = d/c + d = 777/810 = 0·96

Figure 1.1 Approaches to evaluating evidence about diagnosis.

low negative likelihood ratios (LR < 0.1) are generally considered to be clinically useful. The EMA test clearly falls into this category. The authors draws attention to the fact that the probability that a specific patient actually has celiac disease (based on a positive test), or does not have it (based on a negative test), also depends on the pretest odds of the patient having the disease (see Table 1.1).

If the child in question, whose pretest likelihood of celiac disease is estimated to be 8%, has a negative test it may be concluded that the child almost certainly does not have celiac disease; on the other hand, if the child has a positive test, the likelihood of him or her having celiac disease is still only 65%.

As Gregor and Alidina point out, the implications of misdiagnosis must be considered carefully. In the circumstance of a positive test in the child with non-specific symptoms the physician and the child's parents should consider whether it is now reasonable to proceed to intestinal biopsy to confirm the diagnosis, rather than recommending a gluten-free diet, presumably for life. If a search for other clinical or laboratory clues reveals that celiac disease is very likely to be the correct diagnosis, the pretest likelihood may

Table 1.1 The anti-endomysial antibody (EMA) test for celiac disease. Dependence of post-test likelihood of celiac disease on pretest likelihood, assuming positive LR = 31, negative LR = 0.06.

Pretest likelihood of celiac disease	Post-test likelihood with a positive EMA test (%)	Post-test likelihood with a negative EMA test (%)
8% (non-specific symptoms, positive family history)	65	0.5
50% (more specific symptoms)	97	6
0.25% (population screen)	8	0.02

Data from Chapter 10.

be as high as 50%. This would raise the post-test likelihood to 97%. The physician and parents may be comfortable accepting the diagnosis and proceed to a trial of a gluten-free diet, rather than subjecting a young child to intestinal biopsy. This is an excellent example of how a skilled clinician must integrate the principles of evidence-based medicine with traditional clinical skills and judgment.

Offering a prognosis

Example: A 50-year-old woman with recently diagnosed celiac disease. has learned at a meeting of the local celiac society that patients with celiac disease have a substantial increase in the risk of developing a number of cancers and that this cancer risk is reduced by strict adherence to a gluten-free diet.

Chapter 10 describes the types of study which are relevant to determination of prognosis and discusses the strengths and weaknesses of case-control and cohort studies.

Gregor and Alidina point out that certain case-control studies which reported very high mortality and malignancy rates may have been subject to selection bias (inclusion of particularly ill or refractory patients) and measurement bias (patients with abdominal symptoms being more likely to undergo investigations such as small bowel biopsy which may lead to a diagnosis of celiac disease). They refer to a British study in which a cohort of patients with celiac disease was assembled and followed for ten years. This design attempts to minimize the biases that are inherent in the case-control studies. Table 1.2 shows that the risk of certain cancers is increased compared to the risk in the general population. Table 1.3 shows that strict adherence to a gluten-free diet significantly reduced this risk and may have eliminated the excess risk for several of the identified cancers.

Table 1.2 Cancer mortality in 210 patients with celiac disease at the end of 1985.

Site of cancer	ICD8	O	E	O/E	P
All sites	140–208	31	15.48	2.0	[b]
Mouth and pharynx	141–147	3	0.31	9.7	[a]
Esophagus	150	3	0.24	12.3	[a]
Non-Hodgkin's lymphoma	200, 202	9	0.21	42.7	[b]
Gastrointestinal tract	151–154	3	3.07	1.0	NS
Remainder		13	11.65	1.1	NS

[a] $p < 0.01$.
[b] $p < 0.001$.
O: observed numbers; E: expected numbers.
Source: Holmes GKT *et al. Gut* 1989; **30**: 333–338 [25].

On the basis of this evidence it is reasonable to advise the patient that her disease does carry with it an increased risk of certain relatively uncommon cancers and that adherence to a strict gluten-free diet appears to minimize this increased risk.

Recommendations concerning therapy

We have provided examples of how evidence concerning the use of diagnostic tests and prognosis can be analyzed and incorporated into clinical practice. Most chapters in this book deal more extensively with evidence concerning therapy and rely heavily on data from randomized trials and meta-analyses.

Example: Should a 28-year-old woman who has had an uncomplicated resection of the terminal ileum for Crohn's disease receive maintenance therapy with an S-aminosalicylate (ASA) product? Prior to the surgery she had had steroid-dependent disease and had failed treatment with both azathioprine and methotrexate.

A search of the literature for placebo-controlled randomized trials of 5-ASA for maintenance of remission in patients with a surgically induced remission of disease would reveal several trials. The largest published trial is that of McLeod and colleagues, who randomized 163 adult patients to receive either 3 g/day of 5-ASA or a placebo following surgery [26]. The primary outcome of interest was the recurrence of active Crohn's disease as defined by the recurrence of symptoms and the documentation of active disease either radiologically or endoscopically. At

Table 1.3 Cancer morbidity by diet group.

Site of cancer	Diet group[a]	No.	O	E	O/E	P
All sites	1	108	14	9.06	1.5	
	2	102	17	6.42	2.6	[c]
Mouth, pharynx,	1	108	1	0.33	3.0	
esophagus	2	102	5	0.22	22.7	[c]
Non-Hodgkin's	1	108	2	0.12	16.7	[b]
lymphoma	2	102	7	0.09	77.8	[c]
Remainder	1	108	11	8.61	1.3	
	2	102	5	6.11	0.8	

[a] Diet group 1, strict adherence to gluten-free diet; group 2, reduced gluten diet or normal diet. Source: Holmes G KT *et al. Gut* 1989; **30**: 333–338 [25].
[b] $p < 0–01$.
[c] $p < 0.001$.

- **Are the results valid?**

1 Was the assignment of patients to treatment really randomized (and the randomization code concealed)?
2 Were all patients who entered the study accounted for at its conclusion?
3 Were the clinical outcomes measured blindly?

- **Is the therapeutic effect important?**

1 Were both statistical and clinical significance considered?
2 Were all clinically important outcomes reported?

- **What are the results?**

McLeod RS, Wolff BG, Steinhart AH *et al*. Prophylactic mesalamine treatment decreases postoperative recurrence of Crohn's disease. *Gastroenterology* 1995;**109**:404–13.

Randomized controlled trial in which 163 patients with Crohn's disease who had all visible disease resected were randomized to receive mesalamine (Pentasa) 3 g daily or a placebo for a median period of 34 months. Primary outcome was recurrent Crohn's disease defined by recurrence of symptoms and radiographic or endoscopic documentation of recurrence.

	Recurrent Crohn's disease		Risk (%)	ARR (%)	RRR (%)
	Yes	No			
5-ASA	27	60	31	10	24
Placebo	31	45	41	–	–

ARR, absolute risk reduction; RRR, relative risk reduction.

- **Are the results relevant to my patient?**

1 Were the study patients recognizably similar to my own?
2 Is the therapeutic maneuver feasible in my practice?

Figure 1.2 Elements of a valid and useful randomized trial.

the end of the follow-up period (maximum duration 72 months, median duration 34 months), 31% of patients who received active treatment remained in remission compared with 41% of those who received a placebo ($p = 0.031$); 5-ASA was well tolerated. A low proportion of patients developed adverse reactions in the control and active treatment groups. One patient treated with 5-ASA developed pancreatitis that was attributed to the study drug. The results of this study can be evaluated using the guidelines described in Figure 1.2, which is modeled after the approach of Kitching *et al.* [24].

Are the results of this study valid?

A review of the methods section of the article confirms that an appropriate method of randomization was employed (computer-generated in permutated blocks), which insured concealment of the randomization code [26]. Furthermore, inspection of the baseline characteristics of the treatment and control groups shows that they are well balanced with respect to such confounding variables as the time from surgery to randomization. This information further supports the legitimacy of the randomization process. Assessment of the method of randomization is important, because non-randomized designs are especially vulnerable to the effects of bias. Studies which employ "quasi-randomization" schemes such as allocation to treatment according to the day of the week or alphabetically by the patient's surname have been shown to consistently overestimate the treatment effect identified by RCTs that employ a valid randomization scheme [27, 28]. However, it may be noted that 87 patients were randomized to 5-ASA, compared with only 76 patients in the control group. This observation raises the concern that the analysis might not have been done according to the "intent to treat" principle which specifies that patients are analyzed in the group to which they were originally assigned, irrespective of the treatment that was ultimately received. The use of this strategy reduces the possibility of bias, which might occur if investigators selectively withdrew from the analysis patients who had done poorly or experienced toxicity. For this reason, the intent to treat principle yields a conservative estimate of the true benefit of the treatment. However,

detailed review shows that in this study the discrepancy in patient numbers occurred because five patients who were randomized to the active treatment group withdrew consent prior to receiving the study medication and were not included. Thus, it appears that the analysis was based on the intent to treat principle.

Approximately 10% of patients in both treatment groups had incomplete follow-up. Methodologically rigorous studies have a very low proportion of patients for whom data are missing. This issue is important, since patients who are lost to follow-up usually have a different prognosis from those for whom complete information is available. If there is incomplete follow-up data for a substantial proportion of patients then the results are uninterpretable [29].

Turning to an assessment of the outcomes in this study, both the patients and investigators were unaware of the treatment allocation. Blinding is used to reduce bias in the interpretation of outcomes. This is especially important when a subjective outcome is evaluated [30]. In this study, objective demonstration of recurrent disease (endoscopy and/or radiology) was required in addition to the more subjective measure of the introduction of treatment for recurrent symptoms. Thus, the reader can be satisfied that the primary outcome measure was both clinically meaningful and objectively assessed.

Finally, the data analysis and results should be examined. A great deal of useful information can be obtained by reviewing the assumptions that were used in the sample size calculation. In this study, which analyzes a difference in proportions, the investigators had to define four variables: the alpha (type 1) error rate, the beta (type 2) error rate, the expected proportion of patients who would be expected to relapse in the placebo group, and the minimum difference in the rate of relapse which the investigator wished to detect. In this publication these parameters are easily identified. The rate of symptomatic recurrence was estimated to be 12.5% per year and it was anticipated that treatment with 5-ASA would reduce this rate by 50% to an absolute value of 6.25% per year. In contrast to the expected 50% relative risk reduction which was anticipated, the three-year actuarial risk of recurrence was 26% in the treatment group compared to 45% in the group that received 5-ASA ($p = 0.039$). Therefore, the relative risk reduction ((45–26%)/45% = 42%) is slightly lower than the figure which the investigators considered to be clinically meaningful. Furthermore, the probability of a type 1 error is described as a one-tailed value of $p = 0.05$. This implies that one-tailed statistical testing was used to derive the p value of 0.039. The use of one-sided statistical testing raises legitimate concerns regarding the statistical inferences made in the study [31]. It is inappropriate to hypothesize that 5-ASA therapy could only be beneficial, given that the drug can cause diarrhea and colitis [32]. For these reasons, uncertainty exists regarding both the clinical and statistical interpretation of these data.

Are the results of this valid study important?

To assess the importance of this result it is necessary to quantify the magnitude of the treatment effect. How the evidence is presented may influence both physicians and patients in making choices. The most basic means of expressing the magnitude of a treatment of fact is the absolute risk reduction (ARR), which is defined as the proportion of patients in the experimental group with a treatment success minus the proportion of patients with this outcome in the control group. In this instance the annual rate of relapse in the placebo-treated patients was 15% (success rate of 85%) compared with 8.7% (success rate of 91.3%) in those who received the active treatment. This yields an ARR of 6.3%. The number needed to treat (NNT), the number of patients with Crohn's disease who would have to be treated with 3 g/day of 5-ASA to maintain remission over a year, can be calculated as the reciprocal of this number, and is 16. Alternative ways of describing effectiveness include calculating the observed relative risk reduction (RRR = 63/15) of 42%, or even stating that about 90% of patients respond to maintenance therapy, ignoring the substantial placebo effect which is evident. The evidence presented as the ARR or NNT, rather than the numbers which show the treatment in a more favorable light, may still lead the physician to recommend this form of treatment and cause the patient to choose to accept this strategy over no intervention. However, the expectations of the physician and patients are likely to be more realistic than they may be if the physician accepts and promotes in an uncritical way the information that 90% of patients who receive 5-ASA maintenance therapy will remain in remission over one year [33].

Are these results applicable to my patient?

Following an assessment of the validity of the evidence using the criteria described in the preceding paragraphs, it is necessary to decide whether the conclusions of the study are relevant and important to the individual patient. An initial step is to evaluate the demographic characteristics of the patients in the RCT and compare them to those of the patient in question. If the patient for whom maintenance therapy is being considered is similar to the patients who were evaluated in the trial, it is reasonable to assume that she will experience the same benefit of therapy and is at no greater risk for the development of adverse drug

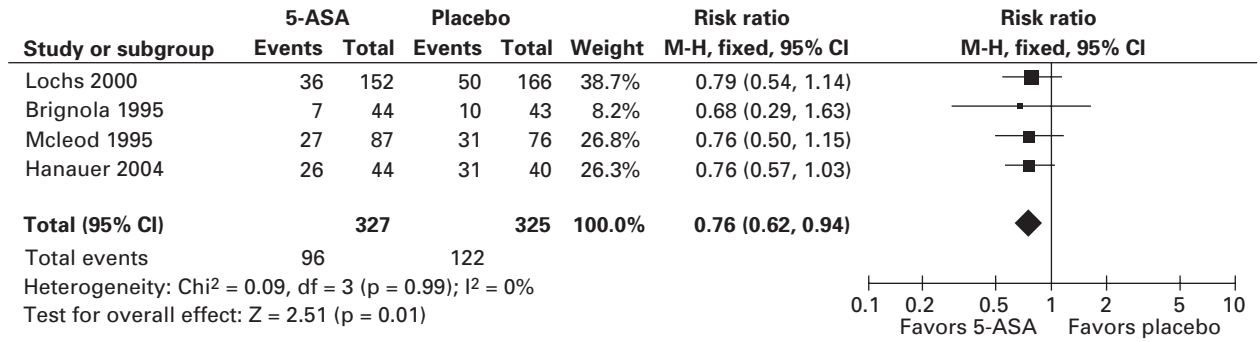

Figure 1.3 Interventions for prevention of post-operative recurrence of Crohn's disease.
Source: Doherty G, Bennett G, Patil S, Cheifetz A, Moss AC. Interventions for prevention of post-operative recurrence of Crohn's disease. In: *Cochrane Database of Systematic Reviews* 2009, Issue 4. Art. No.: CD006873. DOI: 10.1002/14651858.CD006873.pub2.

reactions. Alternatively, this patient may have characteristics that make it unlikely that a benefit from 5-ASA will be realized. For example, if the patient had residual active Crohn's disease it would be difficult to generalize the results of the study of McLeod *et al.*, since the patients in this trial had resection of all visible disease prior to study entry [26].

At this point, if we accept that the results are generalizable to our patient example, the relative risks and benefits of the therapy must be weighed and the patient's preferences should be considered. Evaluation of the data reveals that the trial was methodologically rigorous and evaluated an important outcome. However, it is doubtful whether conventional statistical significance was demonstrated. This raises the question of whether the observed differences between the treatment groups might have occurred by chance. Furthermore, the magnitude of the treatment effect is relatively small. In presenting to the patient the benefit of an annual reduction in the risk of recurrence of 6.3% it is also necessary to consider the cost and inconvenience of taking medication for an asymptomatic condition. One observation in favor of recommending the treatment is that the risk of serious toxicity with 5-ASA appears to be low.

Because there is a degree of uncertainty concerning the true benefit of 5-ASA maintenance therapy based on analysis of this single RCT, it would be prudent to review additional published data. A meta-analysis of 5-ASA therapy has been published [34]. Meta-analysis, the process of combining the results of multiple RCTs using quantitative methods, is an important tool for the practitioner of EBM. Pooling the results of multiple RCTs increases statistical power and thus may resolve the contradictory results of individual studies. Combining data from RCTs statistically also increases the precision of the estimate of a treatment effect. Moreover, the greater statistical power afforded by meta-analysis may allow insight into the benefits of treatment for specific subgroups of patients. These properties are particularly relevant to the case under consideration, given the previously identified concerns.

The meta-analysis summarized data from 15 RCTs which evaluated the efficacy of 5-ASA maintenance therapy in 1371 patients with quiescent Crohn's disease. Patients were randomly assigned to receive either 5-ASA or placebo for treatment periods of 4–48 months. Although 5-ASA was superior to placebo in 13 of the 15 studies, the results of only two trials were statistically significant. Separate analyses were done using data from the four trials that included patients with a surgically induced remission (see Figure 1.3) in distinction to those that evaluated patients after a medically induced remission. Sensitivity analyses assessed the response to therapy in specific subgroups of patients. The overall analysis concluded that 5-ASA has a statistically significant benefit; the risk of symptomatic relapse in patients who received 5-ASA was reduced by 6.3% (95% confidence interval –10.4% to –2.1%, 2 p = 0.0028), which corresponds to an NNT of 16. Importantly, the greatest benefit was observed in the four trials that evaluated patients following a surgical resection. In these studies there was a 13.1% reduction in the risk of a relapse (95% CI: –21.8 – –4.5%, 2 p = 0.0028), which corresponds to an NNT of 8. No statistically significant effect was demonstrable in the analysis, which was restricted to the patients with medically induced remission.

Are the results of this meta-analysis valid and reliable?

Figure 1.4 provides some useful guidelines for the interpretation of overview analyses. It is important that a comprehensive search strategy be adopted since publication bias, the selective publication of studies with positive results, is an important threat to the validity of meta-analysis [35]. This criterion was met. Camma and colleagues' review of the literature was extensive and not limited to English lan-

- Are the results of this overview valid and reliable?

1 Is it an overview of randomized trials of treatments?
2 Does it include a methods section that describes:
  (a) finding and including all the relevant trials?
  (b) assessing their individual validity?
  (c) using valid statistical methods that compare "like with like" stratified by study?
3 Were the results consistent from study to study?
4 Are the conclusions based on sufficiently large amounts of data to exclude a spurious difference (type 1 error) or missing a real difference (type II error).

- Are these applicable to your patient?

Differences between subgroups should only be believed if you can say "yes" to all of the following:

1 Was it hypothesized before the study began (rather than the product of dredging the data), and has it been confirmed in other, independent studies?
2 Was it one of just a few subgroups analyses carried out in this study?
3 Is the difference both clinically (beneficial for some but useless or harmful for others) and statistically significant?
4 Does it really make biologic and clinical sense?

Figure 1.4 Approaches to evaluating evidence concerning overviews.
Reproduced from Yusuf S *et al.*, eds. *Evidence-based Cardiology*. BMJ Books, London, 1998 [22].

guage publications. The investigators also searched review articles, primary studies and abstracts by hand. Quality scores were used to evaluate the validity of the individual studies and a sensitivity analysis was done which assessed the effect of trial quality on the result. No important change in the overall result was noted when studies of lower quality were excluded from consideration. However this type of analysis was not carried out in the analysis of the subgroups of four trials (411 patients) which evaluated 5-ASA after a surgically induced remission.

One of the included studies, that of Caprilli *et al.*, which involved 95 patients, showed a greater benefit for 5-ASA than any other trial, medical or surgical, which has been performed [36]. An important methodological deficiency of this RCT was the failure to conceal the treatment allocation from the investigators. Since these physicians were aware of the treatment assignment, and the definition of relapse used required clinical interpretation, it is possible that the 27% reduction in the risk of relapse identified is an overestimation of the true treatment effect. Accordingly, the inclusion of the results of this study in the subgroup analysis of the surgical studies may overestimate the true benefit of 5-ASA. Furthermore, Camma *et al.* did not include an additional trial by Lochs *et al.*, which was only available as a preliminary report at the time the meta-analysis was done [37]. This study, which is the largest RCT to evaluate 5-ASA following surgery, assigned 318 patients to receive either 4 g of active drug or a placebo for 18 months. Although Camma and colleagues described this study as "confirming" a benefit of 5-ASA after surgery, the results are not impressive. Only a 6.9% reduction in the rate of relapse was observed in patients who received the active treatment (24.5% 5-ASA compared with 31.4% placebo). This difference was *not* statistically significant.

This example underscores the importance of updating systematic reviews as new information becomes available, which is the approach of the Cochrane Collaboration, but not of reviews in conventional publications. When the data provided by Lochs *et al.* were aggregated with those of the other trials, the overall estimate of benefit for 5-ASA was less (ARR 4%, NNT 25) [38]. On the basis of these data it can be concluded that 5-ASA may be an effective maintenance therapy following surgery, but the magnitude of the treatment effect is modest at best.

Chapter 11 includes a meta-analysis performed as part of a Cochrane review that also supports this conclusion (see Figure 1.3).

Are these results applicable to our patient example?

The meta-analysis of surgical trials by Camma *et al.* provides important information to the clinician who must decide whether or not to offer patients 5-ASA for maintenance therapy. The concern regarding statistical significance raised by the critique of the McLeod study has been reduced. It seems likely that the beneficial effect of 5-ASA following surgery is real. However, although the majority of the criteria outlined in Figure 1.4 have been met, the issue of clinical relevance remains. The most optimistic estimate of the size of the treatment effect, derived from the meta-analysis, is an NNT of 8. However, given the possibility of bias in the study of Caprilli *et al.*, a more conservative estimate could be based on the data of Lochs and

colleagues from the single large randomized trial which yielded an NNT of 15, or from the revision by Sutherland of Camma's meta-analysis that yielded an ARR of only 4%, and an NNT of 25.

In presenting this information to the patient the following points should be emphasized.

- The existing data suggest that 5-ASA is not effective, or at the most, very marginally effective.
- The annual risk of relapse following surgery is relatively low without treatment.
- 5-ASA therapy is safe.
- The cost of 5-ASA therapy is approximately US$70 per month.
- To derive a benefit from the treatment the medication must be taken on a regular basis. This requires the patient to take six pills each day.

Patients undoubtedly will react in different ways to this information. Our patient chose not to accept this therapy.

Rationale for a book on evidence-based gastroenterology and hepatology

Gastroenterologists, hepatologists and general surgeons are fortunate to have many excellent textbooks that provide a wealth of information regarding digestive diseases. Such traditional textbooks concentrate on the pathophysiology of disease and are comprehensive in their scope. *Evidence-based Gastroenterology and Hepatology* is not intended to replace these texts, since its focus is on clinical evidence.

Excellent electronic databases are available, and many traditional publications contain relevant research evidence and important summaries and reviews to support evidence-based practice. However, Cumbers and Donald have found that physicians in clinical practice find the acquisition of data from these sources time consuming [39]. Their study revealed that even locating relevant articles required on average three days for practitioners with an on-site library and a week for those without such a facility. This book has been written for the purpose of saving valuable time for busy practitioners of gastroenterology and hepatology, and for general internists and general surgeons who deal with substantial numbers of patients with disorders ranging from gastroesophageal reflux disease to liver transplantation.

It has been extensively revised since the second edition was published in 2004, in order to provide more recent evidence that serves as the basis for recommendations. For example, we present data from Cochrane reviews that summarize the strong evidence that anti-TNF agents are effective for both induction and maintenance of remission of Crohn's disease, along with a careful consideration of the adverse effect profile of these agents.

The book cannot claim to be comprehensive. However, the third edition has been expanded significantly, with new chapters on eosinophilic esophagitis, travelers' diarrhea, antibiotic- associated diarrhea, non-invasive markers for the diagnosis of fibrosis, drug-induced liver injury, liver biopsy, and hepatic outflow syndromes and splanchnic thrombosis. In addition, all chapters have been extensively revised and updated to reflect current evidence. A limitation of any textbook is the timeliness of the information that it is possible to provide in print form. New evidence accumulates rapidly in clinical medicine and it is impossible to include the most up-to-date information in a textbook because of the time required for production. To meet the needs of our readers for the most timely information the editors have endeavored to include, where possible, new evidence that became available during the editorial process. It is also planned to produce electronic updates of chapters at regular intervals. These updates will appear on the Evidence-Based Medicine Series website: http://www.evidencebasedseries.com/Summary of updated evidence in the Third Edition.

These summaries highlight the most significant changes to *Evidence-based Gastroenterology and Hepatology* since the second edition, particularly regarding treatment recommendations in specific conditions. The full discussion of the evidence can be found in the relevant chapters.

Part I: Gastrointestinal disorders

Gastroesophageal reflux disease (Chapter 2)

Data support the use of empiric antisecretory therapy for patients presenting with symptoms thought to be caused by GERD, without performing confirmatory diagnostic testing. PPI are significantly better than H_2-RA for healing esophagitis and relieving symptoms. There are insufficient data to support routinely using PPI doses higher than standard doses for healing esophagitis, treating symptomatic GERD or atypical symptoms of GERD, although higher doses may be effective for preventing relapse that occurs at standard doses. The most cost-effective strategies are PPI based "step-down" or PPI "on-demand" approaches. Laparoscopic fundoplication is an effective alternative to medical therapy, particularly for patients whose symptoms responded to medication.

Barrett's esophagus (Chapter 3)

Aggressive anti-reflux therapy with either high-dose PPI or surgery has not been shown to revert Barrett's esophagus to normal squamous mucosa or reduce the risk of developing cancer. Although endoscopic ablative therapy is a reasonable option in the Barrett's patient with high grade dysplasia or superficial adenocarcinoma, these therapies are not recommended for the Barrett's patient without neoplasia, since continued surveillance will still be required,

complications are frequent and the risk/benefit ratio is not established. Estimates of the cost effectiveness of surveillance of Barrett's esophagus vary widely, and it is not possible currently to make a recommendation for population screening for Barrett's either in the general population or in those with chronic GERD.

Esophageal motility disorders (Chapter 4)
The costs and cost effectiveness of Botulinum toxin injections and pneumatic dilation for achalasia are lower than the cost of Heller myotomy. In the longer term, pneumatic dilatation appears to be more cost-effective than Botulinum toxin injection therapy.

Eosinophilic esophagitis (Chapter 5)
Conventional oral corticosteroids and swallowed inhaled fluticasone both appear to be effective for this condition. However, the adverse effects of oral steroids are more frequent and severe, and inhaled (swallowed) steroids should be used as initial treatment for uncomplicated EE.

Ulcer disease and *Helicobacter pylori* infection (Chapter 6)
Half of ulcer bleeding may be attributable to NSAIDs, and patients who are also positive for *H. pylori* have a synergistically high risk of re-bleeding. *H. pylori* eradication significantly reduces ulcer re-bleeding rates. Clarithromycin resistance accounts for most treatment failures. When a clarithromycin-based eradication regimen has failed, it is not worthwhile to administer it again. More effective options include PPI/amoxicillin/metronidazole or PPI with amoxicillin and levofloxacin for ten days, as well as more conventional bismuth-based quadruple regimens.

NSAID induced gastroduodenal toxicity (Chapter 7)
H. pylori contributes to an excess ulcer-risk in NSAID naive patients, whereas ulcers occurring in long-term NSAID users are probably largely caused by the NSAIDs, irrespective of *H. pylori* status. It is appropriate to eradicate *H. pylori* in NSAID naive patients prior to starting chronic ASA or NSAID therapy. However, *H. pylori* eradication alone appears to be insufficient for ulcer prophylaxis in chronic non-ASA NSAID users. Misoprostol prophylaxis and substitution of COX-2 inhibitors appear to reduce the risk of developing endoscopically diagnosed gastric ulcers by 80% and the risk of complicated ulcers by 50%.

Functional dyspepsia (Chapter 9)
Patients undergoing endoscopy for dyspeptic symptoms tend to be more satisfied and have improved quality of life and subsequently create significantly lower health care costs than patients initially treated empirically. The practical bottom line for use of PPIs in functional dyspepsia is that it is reasonable to give patients a trial of 4–8 weeks of therapy, with the understanding that heartburn is a predictor of response and that the majority of patients will not respond (NNT = 15). *H. pylori* infection is present in 30–70% of patients, and eradication may lead to long-term symptom improvement in a small proportion of these patients. There is no convincing evidence for the use of prokinetic or antidepressant medications.

Celiac disease (Chapter 10)
Although the human recombinant anti-tissue transglutamase antibody test (tTG) has a sensitivity of 96% and a specificity of 99% for diagnosis of celiac disease in some studies, the sensitivity in other studies is considerable lower. The tTG will likely remain as an adjunct to endoscopy for the diagnosis of celiac disease, rather than a replacement. A substantial amount of evidence demonstrates a lack of toxicity of oats in newly diagnosed patients with celiac disease, and in celiac disease patients in remission.

Crohn's disease (Chapter 11)
Infliximab, adalimumab and certolizumab have all been shown to be effective for induction and maintenance of remission. There is emerging evidence that early combined immunosuppression with infliximab combined with azathioprine and, if necessary, steroids is superior to conventional management with corticosteroids, followed in sequence by azathioprine and infliximab in terms of steroid-free remission and avoidance of surgery at 26 and 52 weeks. Serious adverse events are not significantly more frequent in the early combined immunosuppression group. The combination of infliximab and methotrexate, although safe, has not been shown to be more effective than infliximab alone in Crohn's disease patients who are also receiving treatment with prednisone.

Although an earlier small randomized placebo-controlled trial suggested that omega-3 fatty acids are effective for maintenance of remission in patients at a relatively high risk of relapse of Crohn's disease, two large randomized trials that included 762 patients have now shown that this approach is not effective.

Natalizumab (300 mg or 3 to 4 mg/kg) is effective for induction of clinical response and remission in patients with moderately to severely active Crohn's disease. One patient with Crohn's disease treated with natalizumab in combination with azathioprine developed progressive multifocal leukoencephalopathy (PML). A retrospective investigation suggests that the incidence of PML is approximately 1 case per 1000 patients.

Ulcerative colitis (Chapter 12)
Probiotic preparations of a specific lyophilized *E. coli* strain, or mixtures of several bacteria, when added to conventional therapy, may increase rates of induction and enhance maintenance of remission in mild ulcerative colitis. An

apheretic technique that removes granulocytes from the blood of patients is ineffective. Infliximab is effective for induction of remission in severe refractory disease.

Pouchitis (Chapter 13)
Small controlled trials have demonstrated the efficacy of ciprofloxacin and metronidazole for acute pouchitis, of budesonide enemas for active chronic pouchitis, and of probiotic bacteria for maintaining remission of chronic pouchitis and for prophylaxis.

Microscopic colitis (Chapter 14)
Budesonide is effective for short-term treatment, but the optimal strategy for long-term management needs further study. The long-term prognosis is good, and the risk of complications including colonic cancer is low.

Drug-induced diarrhea: (Chapter 15)
Faced with an aging patient population and an ever increasing population of diabetic patients physicians should be aware that cholinesterase inhibitors, increasingly used in Alzheimer's disease, produce diarrhea in 14% of patients and metformin produces malabsorptive diarrhea in up to 50% of patients.

Metabolic bone disease in gastrointestinal disorders (Chapter 16)
Bisphosphonate therapy reduces the risk of vertebral fractures in IBD patients by 6.3%. The improvement in BMD in IBD patients taking infliximab + a bisphosphonate may be greater than that observed with a bisphosphonate alone.

Chronic PPI users may not have an increase in the risk of hip fracture with PPI use if they have no other identifiable risk factor. If there is a true association (as opposed to an association with confounding variables) between PPI use and decreased bone density, the fracture risk in chronic PPI users is low. The increased risk may be clinically relevant in patients with multiple other risk factors for osteoporosis, especially in patients with long-term, high-dose therapy.

Colorectal cancer in ulcerative colitis: surveillance (Chapter 17)
Chromoendoscopy with targeted biopsies may increase the yield of surveillance colonoscopy. Estimates of sensitivity and specificity are made from an increased number of large surveillance programs.

Colon cancer screening (Chapter 18)
The sensitivity of fecal occult blood testing, when three stools are tested using a sensitive immunochemical technique, approaches 90%. The miss rate for significant lesions at colonoscopy may be as high as 5%. There is consistent evidence that colonoscopy is less effective for reduction in right-sided CRC than it is for left-sided CRC. Colonoscopy may be more cost effective than CT colonography, depending on relative procedural costs.

Prevention and treatment of traveler's diarrhea (Chapter 19)
This new chapter summarizes the evidence for a number of interventions for chemoprohylaxis (with Rifaximin being the recommended agent). Early evidence for the efficacy of immunoprohylaxis is also presented.

***Clostridium difficile* disease (Chapter 20)**
Preventive strategies, including enhanced infection control programs and antibiotic stewardship, are effective in reducing the incidence of CDAD. Complicated cases of CDAD have increased in frequency in some outbreaks, and it is postulated that the emergent hypervirulent BI/NAP10/27 strain is responsible. Evidence still favors metronidazole for treatment of initial episodes, but vancomycin appears to be the more effective treatment for severe disease. Observational studies suggest that tapered and pulsed vancomycin regimens may be effective for recurrent disease.

Irritable bowel syndrome (Chapter 21)
Colonic investigation has a very low yield in patients presenting with symptoms that are highly suggestive of IBS in the absence of alarm features. Screening to exclude celiac disease and thyroid dysfunction appears to be of value, but a panel of blood tests, including ESR and CRP, often performed when patients with suspected IBS are first seen in the outpatient clinic, has a low yield in detecting organic disease. Data to support the role of lactose hydrogen breath testing to exclude lactose intolerance are conflicting. Soluble fibre, antispasmodics (particularly hyoscine) and peppermint oil may be of some benefit in treatment of IBS patients, although the evidence is less than convincing. Both TCADs and SSRIs are effective (NNT = 4). Probiotics may also be used as second-line interventions in individuals with particularly troublesome abdominal pain and bloating. Psychological interventions should be reserved for individuals who are unresponsive to, or intolerant of, more conventional therapies, since they are time-consuming and expensive.

Ogilvie's syndrome (Chapter 22)
Evidence is presented that polyethylene glycol may reduce the frequency of recurrent cecal dilatation in patients who had initial resolution with neostigmine or decompression. Polyethylene glycol may also prevent the development of acute pseudoobstruction in patients with multiple organ failure admitted to an ICU.

Gallstone disease (Chapter 23)
The natural history of asymptomatic cholelithiasis in diabetic patients appears to be similar to that in the general population, and preventative surgery should not be recom-

mended routinely. Laparoscopic cholecystectomy appears to be as safe as open cholecystectomy and may provide short-term improvement in quality of life. Acute cholecystitis should be treated with early laparoscopic cholecystectomy. In patients with gallstone pancreatitis preoperative ERCP may increase overall morbidity compared with the approach of cholecystectomy with intraoperative cholangiography. Patients with acute severe gallstone pancreatitis should undergo cholecystectomy following resolution of the acute episode, but during the initial hospital stay. Patients with mild to moderate pancreatitis can be considered for early laparoscopic cholecystectomy. Three approaches to the management of common duct stones (open common bile duct exploration, ERCP and sphincterotomy, and laparoscopic common bile duct exploration) have been compared, and conflicting data may relate to variation in operator expertise. The approach to CBD stones should be individualized and based on the type of expertise available at each institution.

Acute pancreatitis (Chapter 24)
Several recent, well-planned RCTs provided no evidence for benefit from the early use of prophylactic antibiotics in severe pancreatitis, and it is no longer recommended.

Obesity (Chapter 25)
The risk for co-morbidities, including cardiovascular disease, type 2 diabetes, obstructive sleep apnea and certain cancers, is reduced up to 40% by bariatric surgery.

Part II: Liver disease

Hepatitis C (Chapter 26)
There are concise guidelines for treatment of naive patients, based on genotype and viral load. Re-treatment is reviewed in detail as much more evidence is available from randomized studies. In the main, relapsers are worth re-treating, but genotype 1 non-responders may not yield cost-effective benefit.

Hepatitis B (Chapter 27)
New therapies and new evaluation of interferon have revolutionized the management of these patients. The importance of monitoring for viral resistance, and using the correct combination, or sequential use, of agents is outlined.

Alcoholic related liver disease (Chapter 28)
In alcoholic hepatitis there has been a consolidation of the evidence for the use of steroids and evidence for "stopping rules", and validation of indices of non-response. Trials of agents to help abstention are reviewed.

Non-alcoholic fatty liver disease (Chapter 29)
This disease spectrum is far better characterized, including the association with the other factors which make up the metabolic syndrome. Whilst diagnostic methodology has improved, specific therapeutic agents are not available. What has been tried so far is summarized, with an outline of future prospects.

Hemochromatosis (Chapter 30)
Genetic hemochromatosis is underdiagnosed in the general population, and overdiagnosed in patients with secondary iron overload. A mild elevation in ferritin is very common, and may be related to obesity with NAFLD, regular alcohol consumption, or inflammation.

The hepatic iron index reported on histological examination has limited use with the advent of genetic testing. The C282Y homozygote is the classic genetic pattern in >90% of typical cases. With other genetic variants severe iron overload is usually seen in the setting of a concomitant risk factor (alcoholism, viral hepatitis, NAFLD). Several studies have documented reversal of fibrosis following iron depletion therapy.

Wilson's disease (Chapter 31)
Despite isolation of several genes, the diagnosis remains a clinical one in patients presenting with abnormal liver function. Trientene and zinc therapy have further documentation of their efficacy.

Primary biliary cirrhosis (Chapter 32)
Updated information on the use of ursodeoxycholic acid (UDCA), particularly in early PBC, still leave this drug as the one most frequently used for this disease, but with gaps in robust evidence for its efficacy.

Autoimmune hepatitis (Chapter 33)
Classical autoimmune hepatitis is well recognized, but "difficult to treat" cases not responding to standard immunosuppression, and the occurrence of overlap syndromes make some cases difficult to diagnose and to treat. Good outcomes have now been obtained with budenoside, tacrolimus and mycophenalate. New diagnostic algorithms have helped distinguish true overlap syndromes from disease variants.

Primary sclerosing cholangitis (Chapter 34)
It is important to exclude IgG4 associated sclerosing cholangitis in the differential diagnosis. Use of UDCA is under scrutiny, and the use of high dose UDCA is reviewed.

Non-histological assessment of fibrosis (Chapter 35)
This is a new chapter, which has evaluated both serum markers and transient elastography and the comparison with liver biopsy for the assessment of fibrosis.

Portal hypertensive bleeding (Chapter 36)
Evolution of combined therapies using endoscopic and pharmacological ones, has replaced single mode approaches.

The use of primary prophylaxis with non-selective beta-blockers has been extended to grade Child C patients with cirrhosis even if varices are small. Antibiotics are now mandatory in the treatment of acute variceal bleeding as they improve control of bleeding, and mortality over and above specific endoscopic and pharmacological methods.

Hepatic outflow syndromes and splanchnic thrombosis (Chapter 37)
This is a new chapter, updating current management of hepatic outflow obstruction with a defined algorithm and use of anticoagulation for portal venous thrombosis.

Ascites, hepatorenal syndrome and spontaneous bacterial peritonitis (Chapter 38)
This is an updated review of therapy, particularly for hepatorenal syndrome, for which terlipressin and albumin improve renal function in about 30% of patients. Selection for antibiotic prophylaxis in a primary setting has expanded beyond patients with low concentrations of albumin in ascites.

Hepatic encephalopathy: treatment (Chapter 39)
The diagnosis and management of minimal hepatic encephalopathy are integral to this chapter, as this is an evolving area, which does affect quality of life, ability to drive safely and so on.

Hepatocellular carcinoma (Chapter 40)
Combined clinical and imaging characteristics have allowed a much better clinical staging system, which guides therapy. Indications and use of loco-regional therapy, liver resection and transplantation, as well as new agents such as sorafenib have evolved, and a rationale basis for therapy is presented.

Fulminant hepatic failure (Chapter 41)
Supportive management has greatly improved over the past few years. Current best practice is reviewed by new authors, and referral and indications for liver transplantation are fully discussed. The current status of liver support devices is evaluated on the basis of the few controlled trials and observational studies.

Liver transplantation: prevention and treatment of rejection (Chapter 42)
New data from randomized trials has been incorporated against a background of the method of diagnosis of rejection and long-term complications and outcomes, where these have been documented.

Liver transplantation: prevention and treatment of infection (Chapter 43)
New authors have completely revised this chapter, dealing with the major cause of morbidity and mortality within one year of transplantation. Hepatologists who may not be in transplant centers, but who follow up patients who have had liver transplant, need to be aware of infectious complications, their diagnosis and treatment.

Management of HCV infection and liver transplantation (Chapter 44)
This is a major clinical problem, for which only recently there has been data from randomized studies and careful prospective observational studies to give some evidence-based guidance for therapy.

Management of HBV infection and liver transplantation (Chapter 45)
The new therapies for HBV infection have revolutionized the outcome for HBV infected patients who come to liver transplantation. Treatment algorithms are now simplified and the importance for monitoring for viral resistance fully outlined.

Liver biopsy (Chapter 46)
This is a new chapter reviewing the evidence base for diagnosis of liver disease, and staging/grading for chronic viral hepatitis. A full review of transjugular liver biopsy is given.

Drug induced liver injury (DILI) (Chapter 47)
This is a new chapter, including etiopathogenetic mechanisms and practical algorithms to reach a diagnosis for this problem, which is being seen more frequently.

References

1 Ruffin JM, Grizzle JE, Hightower NC, McHardy G, Shull H, Kirsner JB. A co-operative double blind evaluation of gastric "freezing" in the treatment of duodenal ulcer. *N Engl J Med* 1969; **281**: 16–19.
2 Dickson ER, Fleming TR, Wiesner RH *et al.* Trial of penicillamine in advanced primary biliary cirrhosis. *N Engl J Med* 1985; **312**: 1011–1015.
3 A Medical Research Council Investigation. Streptomycin treatment of pulmonary tuberculosis. *BMJ* 1948; **ii**: 770–782.
4 Truelove SC, Witts LJ. Cortisone in ulcerative colitis. Final report on a therapeutic trial. *BMJ* 1955: 1041–1048.
5 Liberati A, Apolone G, Nicolucci A *et al.* The role of attitudes, beliefs, and personal characteristics of Italian physicians in the surgical treatment of early breast cancer. *Am J Public Health* 1990; **81**: 38–41.
6 Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. *JAMA* 1992; **268**: 240–248.
7 Mulrow CD. The medical review article: state of the science. *Ann Intern Med* 1987; **106**: 485–488.
8 Jefferson T. What are the benefits of editorials and nonsystematic reviews? *BMJ* 1999; **318**: 135.

9 Cochrane AL. Archie Cochrane in his own words. Selections arranged from his 1972 introduction to "Effectiveness and efficiency: random reflections on the health services" 1972. *Control Clin Trials* 1989; **10**: 428–433.

10 The Cochrane Collaboration. *Cochrane Library* 1999: www.cochrane.org.

11 Sackett DL. Clinical epidemiology. *Am J Epidemiol* 1969; **89**: 125–128.

12 Sackett DL. Interpretation of diagnostic data: 1. How to do it with pictures. *Can Med Assoc J* 1983; **129**: 429–432.

13 Sackett D, Haynes RB, Tugwell P, Guyatt GH. *Clinical Epidemiology: a Basic Science for Clinical Medicine*, 2nd edn. Little, Brown and Company, Boston, MA, 1991.

14 Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research. getting better but still not good. *JAMA* 1995; **274**: 645–651.

15 Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med* 1978; **299**: 926–930.

16 Chalmers TC. Randomization of the first patient. *Med Clin North Am* 1975; **59**: 1035–1038.

17 Resnick RH, Iber FL, Ishihara AM, Chalmers TC, Zimmerman H. A controlled study of the therapeutic portacaval shunt. *Gastroenterology* 1974; **67**: 843–587.

18 Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. *N Engl J Med* 1987; **316**: 450–455.

19 Kernick D. Lies, damned lies, and evidence-based medicine. Jabs and jibes. *Lancet* 1998; **351**: 1824.

20 Scholefield JH, Johnson AG, Shorthouse AJ. Current surgical practice in screening for colorectal cancer based on family history criteria. *Br J Surg* 1998; **85**: 1543–1546.

21 Gabriel SE, Jaakkimainen L, Bombardier C. Risk for serious gastrointestinal complications related to use of nonsteroidal anti-inflammatory drugs. A meta-analysis. *Ann Intern Med* 1991; **115**: 787–796.

22 Yusuf S, Cairns JA, Camm AJ, Fallen EL, Gersh BJ. *Evidence-based Cardiology*, 2nd edn. BMJ Books, London, 2003.

23 Cataldo F, Ventura A, Lazzari R. Anti-endomysium antibodies and celiac disease: solved and unsolved questions. An Italian multicentre study. *Acta Paediatr* 1995; **84**: 1125–1131.

24 Kitching A, Sackett D, Yusuf S. Approaches to evaluating evidence. *Evidence-based Cardiology*. BMJ Books, London, 1998.

25 Holmes GKT, Prior R, Lane MR *et al.* Malignancy in celiac disease: effect of a gluten-free diet. *Gut* 1989; **30**: 333–338.

26 McLeod RS, Wolff BG, Steinhart AH *et al.* Prophylactic mesalamine treatment decreases postoperative recurrence of Crohn's disease. *Gastroenterology* 1995; **109**: 404–413.

27 Chalmers TC, Celano P, Sacks HS, Smith H Jr. Bias in treatment assignment in controlled clinical trials. *N Engl J Med* 1983; **309**: 1358–1361.

28 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; **273**: 408–412.

29 ICH Steering Committee. *ICH Harmonised Tripartite Guideline. Statistical Principles for Clinical Trials*. Section 5.3-Missing Values and Outliers. Geneva: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 1998.

30 Feagan BG, McDonald JWD, Koval JJ. Therapeutics and inflammatory bowel disease: a guide to the interpretation of randomized controlled trials. *Gastroenterology* 1996; **110**: 275–823.

31 Koch GG. One-sided and two-sided tests and p values. *J Biopharm Stat* 1991; **1**: 161–170.

32 Kapur KC, Williams GT, Allison MC. Mesalazine induced exacerbation of ulcerative colitis. *Gut* 1995; **37**: 838–839.

33 Naylor CD, Chen E, Strauss B. Measured enthusiasm: does the method of reporting trial results alter perceptions of therapeutic effectiveness? *Ann Intern Med* 1992; **117**: 916–921.

34 Camma C, Giunta M, Rosselli M, Cottone M. Mesalamine in the maintenance treatment of Crohn's disease: a meta-analysis adjusted for confounding variables. *Gastroenterology* 1997; **113**: 1465–1473.

35 Oxman AD, Cook DJ, Guyatt GH. User's guides to the medical literature. VI How to use an overview. Evidence-Based Medicine Working Group. *JAMA* 1994; **272**: 1367–1371.

36 Caprilli R, Andreoli A, Capurso L *et al.* Oral mesalazine (5-aminosalicylic acid; asacol) for the prevention of postoperative recurrence of Crohn's disease. *Aliment Pharmacol Ther* 1994; **8**: 35–43.

37 Lochs H, Mayer M, Fleig WE *et al.* Prophylaxis of postoperative relapse in Crohn's disease with mesalazine (Pentasa) in comparison to placebo. *Gastroenterology* 2000; **119**: 264–273.

38 Sutherland LR. Mesalamine for the prevention of postoperative recurrence: is nearly there the same as being there? *Gastroenterology* 2000; **118**: 264–273.

39 Cumbers B, Donald A. Evidence-based practice. Data day. *Health Serv J* 1999; **109**: 30–31.