Chapter 7

Evidence-based medicine

Evidence-based medicine (EBM) and clinical pharmacy

EBM has become standard practice during recent years, although it is probably more widely practised in primary care in the UK. The following definition of EBM can be adapted for clinical pharmacy.

Definition of EBM

> EBM is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients.[1]

The authors of the definition go on to state that the practice of EBM requires the integration of individual clinical expertise with the best available external clinical evidence from systematic research.

The second definition comes from the McMaster University website:

> EBM is an approach to healthcare that promotes the collection, interpretation and integration of valid, important and applicable patient-reported, clinician-observed and research-derived evidence. The best available evidence, moderated by patient circumstances and preferences, is applied to improve the quality of clinical judgements.[2]

Evidence-based clinical pharmacy

Borrowing the Sackett definition, a definition might be as follows: *Evidence-based clinical pharmacy is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients.*

This entirely fits with the concept of pharmaceutical care (see 🕮 p.244) and challenges clinical pharmacists not only to keep abreast of developments in their chosen specialty, but also to apply clinical developments to patient circumstances and preferences.

One of Bandolier's maxims is that EBM is essentially 'tools not rules'.[3] Pharmacists need to remember this when applying current best evidence to patient care.

Strengths of evidence

A hierarchy of evidence (Table 7.1) is helpful in avoiding types of studies that are inherently biased. A number of grading systems are currently available which are useful in terms of identifying the level of evidence available and as a tool for categorizing recommendations made in clinical guidelines, for example. For updated information on this topic, see the GRADE website.[4]

Some evidence tables regard large randomized trials as level I evidence. Evidence from levels IV and V should not be overlooked if it is all that is available. Conversely, recommendations should not be made on level V evidence if level I or II evidence is available.

1 Sackett DL *et al.* (1996). *British medical Journal* **312**: 71–2.
2 McMaster University. http://hiru.mcmaster.ca.
3 www.ebandolier.com.
4 www.gradeworkinggroup.org.

Table 7.1 Type and strength of efficacy evidence

I	Strong evidence from at least one systematic review of multiple well-designed randomized controlled trials
II	Strong evidence from at least one properly designed randomized controlled trial of appropriate size
III	Evidence from well-designed trials without randomization, single group, cohort, time series, or matched case-controlled studies
IV	Evidence from well-designed non-experimental studies from more than one centre or research group
V	Opinions of respected authorities, based on clinical evidence, descriptive studies, or reports of expert committees

Further reading

Useful resource for pharmacy is available on the Bandolier website www.medicine.ox.ac.uk/bandolier/booth/booths/pharmacy.html.

Statistical versus clinical significance

Simply because a study finding is statistically significant, does not mean that the finding is important. Large trials or large meta-analyses have the potential to find very small statistically significant differences between groups. An important consideration when interpreting significant findings is assessment of how clinically significant the finding is.

'*Clinical significance*' refers to a value judgement people must make when determining the meaningfulness of the magnitude of an intervention effect.

For example, if an expensive medication was found to significantly ↓ systolic blood pressure (SBP) by an average of 2mmHg, it would be important to consider the clinical merit of the intervention. Would there be any important health benefits to a patient of a ↓ in SBP of just 2mmHg? Would it be worth investing in an expensive intervention if it delivered such a meagre ↓ in SBP? Are there any cheaper medications available that produce greater ↓ in BP?

Well-conducted rigorous randomized controlled trials should recruit enough participants to detect a difference between groups which is determined as clinically significant before the study.

This page intentionally left blank

Odds ratios and relative risk

What is an odds ratio?

The number needed to treat (NNT) is a very useful way of describing the benefits (or harms) of treatments, both in individual trials and in systematic reviews. Few papers report results using this easily interpretable measure. However, NNT calculations come second to working out whether an effect of treatment in one group of patients is different from that found in the control groups. Many studies, particularly systematic reviews, report their results as odds ratios or as a ↓ in odds ratios, and some trials do the same. Odds ratios are also commonly used in epidemiological studies to describe the probable harm an exposure might cause.

Calculating the odds

The odds of an event occurring are calculated as the number of events divided by the number of non-events. For example, 24 pharmacists are on call in a major city. Six pharmacists are called. The odds of being called are 6 divided by 18 (the number who were not called) or 0.33. An odds ratio is calculated by dividing the odds in the treated or exposed group by the odds in the control group. In general, epidemiological studies try to identify factors that cause harm—those with odds ratios >1. For example, if we look at case–control studies investigating the potential harm of giving high doses of calcium-channel blockers to treat hypertension. Clinical trials typically look for treatments that ↓ event rates, and that have odds ratios <1. In these cases, a percentage ↓ in the odds ratio is often quoted instead of the odds ratio. For example, the ISIS-4 trial reported a 7% ↓ in the odds of mortality with captopril treatment, rather than reporting an odds ratio of 0.93.

Relative risks

Few people have a natural ability to interpret event rates that are reported in terms of odds ratios. Understanding risks and relative risks seems to be easier to grasp.

The risk (or probability) of being called in the example already described in 'Calculating the odds' is 6 divided by 24 (the total number on call) or 0.25 (25%). The relative risk is also known as the 'risk ratio', and if reporting positive outcomes, such as improvement, it can be called 'relative benefit'.

Risks and odds

In many situations in medicine, we can get a long way in interpreting odds ratios by pretending that they are relative risks. When events are rare, risks and odds are very similar. For example, in the ISIS-4 study 2231 out of 29 022 patients in the control group died within 35 days: a risk of 0.077 (2231/29 022) or an odds of 0.083 (2231/(29 022–2231)). This is an absolute difference of 6 in 1000 or a relative error of ~7%.This close approximation holds true when we talk about odds ratios and relative risks, provided that the events are rare.

Why use an odds ratio rather than relative risk?

If odds ratios are difficult to interpret, why don't we always use relative risks instead? There are several reasons for continuing with odds ratios, most of which relate to the superior mathematical properties of odds ratios. Odds ratios can always take values between zero and infinity, which is not the case for relative risks.

The range that relative risk can take depends on the baseline event rate. This could obviously cause problems if we were performing a meta-analysis of relative risks in trials with greatly different event rates. Odds ratios also possess a symmetrical property: if you reverse the outcomes in the analysis and look at good outcomes rather than bad outcomes, the relationships have reciprocal odds ratios. Again, this is not true for relative risks.

Odds ratios are always used in case–control studies where disease prevalence is not known: the apparent prevalence depends solely on the ratio of sampling cases to controls, which is totally artificial. To use an effect measure that is altered by prevalence in these circumstances would obviously be wrong, so odds ratios are the ideal choice. This, in fact, provides the historical link with their use in meta-analyses: the statistical methods that are routinely used are based on methods first published in the 1950s for the analysis of stratified case–control studies. Meta-analytical methods that combine relative risks and absolute risk reductions are now available, but more caution is required in their application, especially when there are large variations in baseline event rates.

A fourth point of convenience occurs if it is necessary to make adjustments for confounding factors using multiple regression. When measuring event rates, the correct approach is to use logistic regression models that work in terms of odds and report effects as odds ratios. All of which makes odds ratios likely to be in use for some time—so it is important to understand how to use them. Of course, it is also important to consider the statistical significance of an effect in addition to its size: as with relative risks, it is easy to spot statistically significant odds ratios by noting whether their 95% confidence intervals do not include 1, which is analogous to a <1 in 20 chance (or a probability of <0.05 or gambling odds of better than 19:1) that the reported effect is solely due to chance.

Formula to calculate an odds ratio

$$\text{Odds ratio} = \frac{\text{odds on treatment}}{\text{odds on control}}$$

Where odds ratio = 1, this implies no difference in effect

Formula to calculate a relative risk

$$\text{Risk ratio} = \frac{\text{risk on treatment}}{\text{risk on control}}$$

Where risk ratio = 1, this implies no difference in effect

Binary and continuous data

Broadly, statistical tests can be grouped into those used to compare *binary* (also called 'dichotomous') outcome data and those used to compare *continuous* outcome data. Binary outcomes are those that can only take two possible values, such as dead or alive, pain or no pain, and smoker or non-smoker. Statistical tests on binary data, such as relative risks, compare the rate of an event between the groups; it also makes the calculation of NNT possible. Continuous outcomes are derived from data that can take any value on a scale. Some examples of continuous data include height, BP, time, or the score in a test. Statistical tests on continuous data (e.g. *t* tests) compare the difference between means of each group (see 📖 p.134).

L'Abbé plots

L'Abbé plots are named after a paper by Kristen L'Abbé and colleagues and are an extremely valuable contribution to understanding systematic reviews. The authors suggest a simple graphical representation of the information from trials. Each point on a L'Abbé scatter plot represents one trial in the review. They are a simple and effective way to present a series of results, without complex statistics. The proportion of patients achieving the outcome with the experimental intervention is plotted against the event rate in the control group. Even if a review does not show the data in this way, it is relatively simple to determine this if the information is available.

For treatment, trials in which the experimental intervention was better than the control are in the upper-left section of the plot, between the *y*-axis and the line of equality. If the experimental intervention was no better than the control, the point falls on the line of equality, and if the control was better than the experimental intervention, the point is in the lower-right section of the plot, between the *x*-axis and the line of equality (Fig. 7.1).

For prophylaxis, this pattern is reversed. Because prophylaxis ↓ the number of bad events (e.g. death after myocardial infarction following the use of aspirin), we expect a smaller proportion of patients harmed by treatment than in the control group. So if the experimental intervention is better than the control, the trial results should be between the *x*-axis and the line of equality.

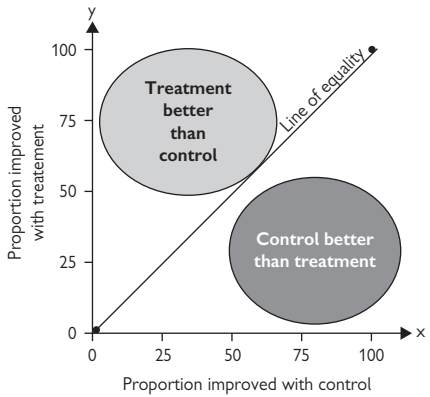

Fig. 7.1 L'Abbé plot for treatment.

Mean difference and standardized mean difference

Analyses of continuous data often show the difference between the means of the groups being compared. In a meta-analysis, this can involve either comparing the mean difference of trials in two groups directly if the unit of measurement of the outcome is the same (e.g. if height is the outcome of interest and all trials measure height in centimetres) or standardizing the outcome measure and comparing the difference between the standardized means if different assessment scales are used to measure subjective conditions, such as mood, depression, or pain.

In a meta-analysis of continuous data, if an experimental intervention has an identical effect as a control (or comparison), the mean difference or standardized mean difference is zero. Therefore if the lower limit of a confidence interval around a mean difference or standardized mean difference is >0, the mean of the experimental intervention group is significantly greater than that of the control group. Similarly, if the upper limit of the confidence interval is <0, the mean of the experimental intervention is significantly lower than that of the control. However, if the confidence interval incorporates the value 0, there is no significant difference between the means of the groups being compared.

Consider the output from a Cochrane review which compared the effect of very low calorie diets (VLCDs) with other interventions for weight loss in patients with type 2 diabetes mellitus (Fig. 7.2). In this case, weight loss is measured in kilograms so there is no need for standardization. As can be seen, the meta-analysis of the two trials indicated that the mean difference in weight between the management with a VLCD and other interventions is –2.95kg. This suggests that patients with type 2 diabetes mellitus on a VLCD are, on average, 2.95kg lighter than patients with type 2 diabetes mellitus on the comparison interventions. However, the range of the 95% confidence intervals includes 0, which indicates that the difference in weight loss between the two groups is not statistically significant.

Review: Long-term non-pharmacological weight loss interventions for adults with type 2 diabetes mellitus
Comparison: 01 VLCD vs different intervention (1–10: fixed models. 11–20: random models, rho = 0.75)
Outcome: 01 weight loss (kg)

Study	Treatment N	Treatment Mean (SD)	Control N	Control Mean (SD)	Weighted mean difference (fixed) 95% CI	Weight (%)	Weighted mean difference (fixed) 95% CI
Wing, 1991a	17	–8.60 (9.20)	16	–6.80 (6.90)		39.5	–1.80 [–7.33, 3.73]
Wing, 1994	48	–14.20 (10.30)	45	–10.50 (11.60)		60.5	–3.70 [–8.17, 0.77]
Total (95% CI)	65		61			100.0	–2.95 [–6.42, 0.53]

Test for heterogeneity chi-square = 0.27 df = 1p = 0.60 I^2 = 0.0%
Test for overall effect z = 1.66 p = 0.1

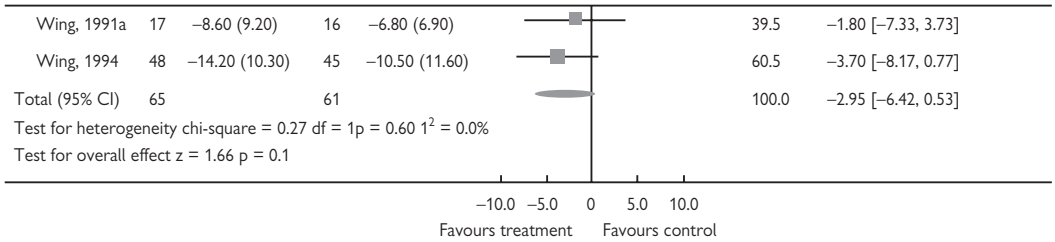

Fig. 7.2 Meta-analysis of a VLCD versus other interventions for weight loss in patients with type 2 diabetes mellitus.

Assessing the quality of randomized studies

Assessment tools for randomized studies are widely available and all have problems because they do not cover all the issues that could be considered to be important. This simple method picks up on the main issues of randomization, blinding, and patient withdrawal from studies (Table 7.2). The maximum quality score is 5 if all the criteria are fulfilled.

In addition, a more general appraisal tool is presented (Table 7.3). It picks up details from the scoring system described in Table 7.2.

Table 7.2 Simple assessment tool for a randomized trial[1]

Is the study randomized?	Score
Yes	1
Is the randomization appropriate?	
Yes—e.g. random number tables	1
No—e.g. alternate patients, date of birth, or hospital number	–1
Was the study double blind?	
Yes	1
Was blinding correctly carried out?	
Yes—e.g. double dummy	1
No—e.g. treatments did not look identical	–1
Were withdrawals and drop-outs described?	
Yes	1

[1] Jadad A *et al.* (1996). Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clinical Trials*. **17**: 1–12.

Table 7.3 General appraisal tool for a randomized trial

Was the method of randomization appropriate (e.g. computer generated)?
Was the study described as 'double-blind'? And was the method of blinding adequate (e.g. double dummy, or identical tablets)?
Was the trial sensitive, i.e. able to detect a difference between treatment groups (e.g. use of a placebo, or additional active groups)?
Were baseline values for each treatment group adequate for trialists to measure a change following treatment?
Were the groups similar at the start of the trial?
Similar patients?
Diagnostic criteria clearly stated?
Similar baseline measures?
Was the size of the trial adequate?
How many patients were there in each group?
Were outcomes clearly defined and measured appropriately?
Were they clinically meaningful?
Were they primary/surrogate outcomes?
Were the outcome data presented clearly?
If multiple tests were conducted, were single positive results inappropriately presented?
Quality score[1]

Randomization	Double-blinding	Withdrawals/ drop-outs	Total score

Critical appraisal of systematic reviews

Systematic reviews are considered to be the best level of evidence if they are well conducted and evaluate a number of randomized trials. They can be particularly useful when seeking to answer clinical questions. However, they are only reliable if the process of the review has followed rigorous scientific principles. Authors should explicitly state the topic being reviewed and have made a reasonable attempt to identify all the relevant studies. The 10 questions listed in Table 7.4 help in that assessment. If the paper fails either of the first two questions, it is not worth proceeding further.

Table 7.4 Ten questions to make sense of a review[1]

For each question answer : Yes, No, or Don't Know
A. Are the results of the review valid?
1. Did the review address a clearly focused issue (e.g. the population, intervention, and/or outcomes)?
2. Did the authors look for the appropriate sort of papers? Check that the authors looked for randomized controlled trials or had clear reasons for including other types of studies.
Is it worth continuing?
3. Do you think the relevant important studies were included? Look for search methods, use reference list, unpublished studies and non-English language articles.
4. Did the authors do enough to assess the quality of the studies included? This would routinely be in the form of an assessment tool for randomized controlled trials.
5. If the results of studies were combined, was it reasonable to do so?
B. What are the results?
6. What is the overall result of the review? Is there a clear numerical expression?
7. How precise are the results? What were the confidence intervals?
C. Will the results help my local situation?
8. Can the results be applied locally?
9. Were all important outcomes considered?
10. Are the benefits worth the harms and costs?

[1] Oxman AD *et al.* (1994). Users guide to the medical literature. VI: How to use an overview. *Journal of the American Medical Association* **272**: 1367–71.

This page intentionally left blank

Critical assessment of papers

When reading a clinical trial paper, it is too easy to read the abstract quickly and skim through the main text. Taking the time to critically evaluate the paper might seem daunting and too time-consuming. In many situations a quick read through is all that is needed. However, if the information gleaned from the paper is going to be used to decide on treatment options or might be used to support a formulary application, a more thoughtful approach is required. The information in this section specifically relates to critically evaluating a clinical trial paper, but the same process, adapted to the content, can be used for other types of clinical paper.

It is not necessary to be a statistician or an expert in trial design to critically evaluate a paper. Much of the evaluation is common sense. A full critical evaluation should take all the following points into account, but even simply bearing them in mind will help you get more out of any paper you read.

- Title—does this accurately reflect the content of the paper? Ideally, the title should state the question under investigation, rather than potentially biasing readers by declaring the results. Cryptic titles are a popular way of attracting readers' attention, but if it is too obscure, could it be because that the authors don't really know what they are writing about? Before progressing, consider how useful this trial is in the clinical setting. If it is too esoteric, it might not be worth reading any further!
- Authors—should be from professions/institutions appropriate to the subject studied. Be cautious with papers authored by pharmaceutical industry employees, but don't dismiss these out of hand. Too many authors might mean that the work is scrappy. Multicentre studies should list the key authors and acknowledge other participants at the end of the paper. Is a statistician listed as an author or acknowledged? This should provide reassurance that the statistics are correct.
- Journal—don't assume that because a paper is published in a mainstream journal it is a good paper. However, be more cautious about papers from obscure journals.
- The introduction—should give relevant background information, building logically to the study topic. If the introduction is waffly or irrelevant, ask yourself if the authors really know what they are writing about.
- Method—a well-written method should give sufficient information for another person to reproduce the study. The information given should include the following.
  - Type of study (e.g. randomized controlled trial, cohort, or case study).
  - Numbers involved, ideally including details of powering.
  - Patient selection and randomization—details of patient demo-graphics should be given and the baseline characteristics of each group should be roughly the same (and should be acknowledged if not).
  - Inclusion/exclusion criteria—consider whether these are appropriate. If there are too many exclusion criteria, the study might not be relevant to the clinical setting.

- Outcome measurements—by now, the question that the authors are trying to answer should be clear. The factors used to measure the outcome should be appropriate and, if possible, directly related to the question. Be cautious of surrogate markers. In many clinical settings, it might be unethical, too invasive, or take too long to use the target outcome. However, check that the surrogate marker closely reflects the target outcome as a whole and not just one aspect of it.
- An appropriate comparator drug should be used at its standard dose. Any new drug should be tested against standard therapy. If a drug is compared with placebo or an outdated or rarely used drug, ask yourself why. With the exception of the study treatment, all other interventions should be the same.
- A randomized controlled trial should ideally be double-blinded (i.e. neither the study participants nor the investigators know which subjects are receiving the study drug and which subjects are receiving the comparator). Sometimes this is not feasible or ethical, but there might be bias if the trial is open-label (both subjects and investigator know who is receiving which treatment) or single-blind (the investigator but not the participants know who is receiving each treatment).
- Be cautious with crossover trials—if the disease studied could improve with time without treatment (especially if it is self-limiting or seasonal), a crossover trial is inappropriate. An adequate 'washout' period between treatments is essential.
- The details of statistical tests should be given—the tests should be appropriate to the type of data presented. Beware of trials that use numerous statistical tests. Why are so many tests needed? Is it that there is nothing to prove? Further discussion of statistical tests is beyond the scope of this topic. Consult relevant textbooks for further information.

- Results—should answer the question originally asked and be easy to comprehend.
  - Graphs and tables should be relevant and clear. Too many graphs and tables suggest that the authors are having difficulty proving their point! Watch labelling of axes on graphs. Sometimes labelling is skewed (e.g. does not start at zero) to give more impressive results.
  - If means are quoted, the variance and/or median should also be quoted. This helps determine whether the mean is a true 'average' or whether extreme values have skewed the results.
  - The results might be statistically significant, but are they clinically significant? Results presented as odds ratios, relative risks, or NNT are generally easier to apply to a clinical setting.
- The discussion—should logically build from the results to answer the original question, one way or another. If the authors make statements such as 'further study is required . . .', ask yourself why. Is this because the original study design was unsuitable? Any doubts or inconsistencies should be dealt with satisfactorily, not just explained away.

- The conclusion—should be appropriate to the data presented and give a definite final answer. If the conclusion is woolly, was there any point in the study in the first place or were the authors just 'paper chasing'?
- The bibliography—should be up to date and relevant. Beware of too many references from obscure journals. You should be able to satisfactorily follow up statements made in the rest of the paper by reference to the original papers quoted.
- Acknowledgements—look for any specialists not in the author list, which might provide reassurance if you had any doubts about the authors' expertise in any angle of the study. Watch out for funding or sponsorship from parties with a vested interest in the outcome of the study (notably the pharmaceutical industry!). However, don't dismiss studies sponsored by the pharmaceutical industry out of hand. Much good work is supported by the pharmaceutical industry.

Further reading

Sackett DL *et al.* (2005). *Evidence-based Medicine*. Churchill Livingstone.
Jones C (2002). Evidence-based medicine. 1: Research methods. *Pharmaceutical Journal* **268**: 875–7.
www.clinicalevidence.com

Guidelines

Guideline development is a common way of either seeking to introduce new practices or attempting to stop some current practices. Guidelines can be time consuming and costly to develop. There is evidence that they can be effective if carefully prepared and peer reviewed. Shekelle *et al.*[1] proposed the following key steps that need to be followed.

- Identify and refine the subject area.
- Create a guideline development group.
- Based on systematic reviews:
  - assess the evidence about the clinical question or condition
  - translate the evidence into a recommendation within the guideline.
- Ensure that the guideline is externally reviewed.

A useful checklist for guidelines is provided by Shaneyfelt *et al.*[2] This review of some 270 guidelines lists some 25 points to consider when preparing a guideline. These include stating the purpose of the guideline, using an expiry date, and grading the recommendations according to the strength of the evidence.

1 Shekelle PG *et al.* (1999). Developing clinical guidelines. *Western Journal of Medicine* **170**: 348–51.

2 Shaneyfelt TM *et al* (1999). Are guidelines following guidelines? *Journal of the Amereican Medical Association* **281**: 1900–5.

Number needed to treat (NNT)

The NNT is a measure of clinical significance and changes view from 'Does a treatment work?' to 'How well does a treatment work?'. This concept is widely used and useful not only in its own right, but also to enable direct comparisons of treatments. The league table of treatments from the Oxford Pain Research Unit (Fig. 7.3) illustrates the value of such an approach. Ideally, we would want an NNT of 1. Although there are treatments that meet this criterion (e.g. anaesthetic agents) in practice NNTs are >1 for the reasons discussed here.

The NNT is defined as follows: the number of people who must be treated for one patient to benefit. The NNT is expressed in terms of a specific clinical outcome and should be shown with confidence intervals.

Calculating the NNT for active treatments

The NNT calculation is based from the understanding of risk ratios (Fig. 7.4). Although the NNT is the reciprocal of the absolute risk reduction, it is not necessary to understand this concept to calculate the NNT. A worked example is included so that the process is transparent. The equation is quite simple, and it is easy to calculate the NNT in published trials using a pocket calculator.

The NNT was initially used to describe prophylactic interventions. The NNT for prophylaxis is given by the following equation:

> 1/(proportion of patients benefiting from the control intervention minus the proportion of patients benefiting from the experimental intervention).

The NNT for active treatment is given by the following equation:

> 1/(proportion of patients benefiting from the experimental intervention minus the proportion of patients benefiting from the control intervention).

From the equation in Fig. 7.4 it should be apparent that any response in the control arm leads to NNT >1. People often ask what a good NNT is; it depends whether the NNT is for treatment (ideally in the range 2–4) or prophylaxis (the NNT is generally larger). Issues such as toxicity have an influence, including the cost. For example, a cheap and safe intervention that prevents a serious disease but has an NNT of 100 might well be acceptable.

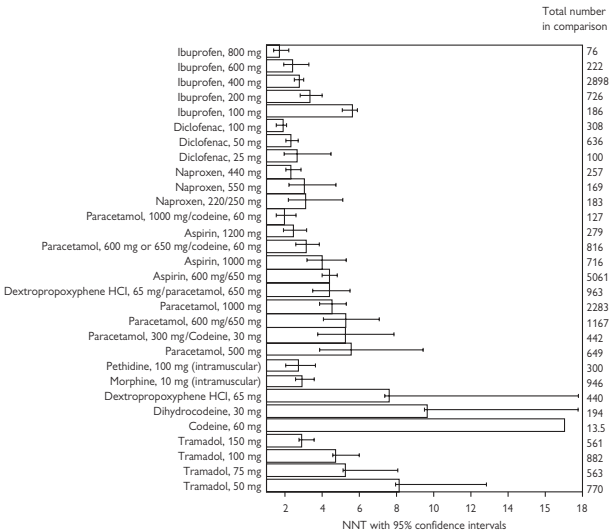

Fig. 7.3 League table of NNT to produce ≥50% pain relief for 4–6h compared with placebo in patients with pain of moderate or severe intensity.

	Controls	Active treatment
Number of patients	N_{con}	N_{act}
Improved = clinical end point	Imp_{con}	Imp_{act}

$$NNT = \frac{1}{\frac{Imp_{act}}{N_{act}} - \frac{Imp_{con}}{N_{con}}}$$

Fig. 7.4 Number needed to treat (NNT).

Using the NNT to express harm

The number needed to harm (NNH) can also be helpful, in addition to the NNT. The NNH is calculated using a similar formula derived from data for adverse events rather than desired effect (Fig. 7.5).

	Controls	Active treatment
Number of patients	N_{con}	N_{act}
AE-number with the adverse	AE_{con}	AE_{act}

$$NNH = \frac{1}{\frac{AE_{act}}{N_{act}} - \frac{AE_{con}}{N_{con}}}$$

Fig. 7.5 Number needed to harm (NNH).

Confidence intervals

Most pharmacists are aware of *p* values in terms of an answer being significant (in a statistical sense) or not. However, the use of *p* is increasingly redundant, and new methods of reporting significance have emerged.

The most common method is the confidence interval, which enables us to estimate the margin of error. For example, if we measured BP in 100 adults, we could derive a mean result. If we then took a further 100 adults and repeated the experiment, we would arrive at a similar, but not identical, figure. The confidence interval, expressed as a percentage, enables calculation of the margin of error and tells us how good our mean is. Generally, the figure is set at 95%, so we can be confident that the true mean lies somewhere between the upper and lower estimates (Fig. 7.6). Expressed a different way, there is only a 5% chance of the result being outside the calculated limits.

The statistics involved are derived from a range of 1.96 standard deviations above and below the point estimated. For a 99% confidence interval, a figure of 2.58 standard deviations is used.

Calculating confidence intervals

Although the formulae are available in standard statistics works, there are a number of confidence interval calculators on the web that require the use of the calculated point estimate and the number of samples to derive the confidence interval at a given percentage.

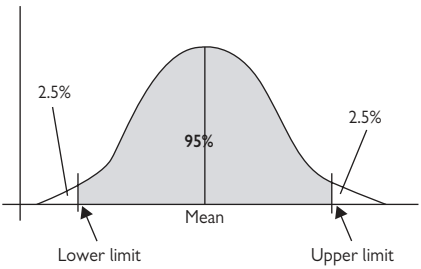

Fig. 7.6 Illustration of the data incorporated within a 95% confidence interval.

This page intentionally left blank

Chapter 8

Herbal medicines

Herbal drugs

The efficacy and safety of herbal drugs present a number of issues to pharmacists. Herbal drugs are more often complex mixtures of active constituents that vary in quality for a number of reasons, such as environmental and genetic factors. Furthermore, the constituents responsible for the claimed therapeutic effects are frequently unknown or only partly explained.

The position is further complicated by the traditional practice of using combinations of herbal drugs, and it is not uncommon to have as many as five or more herbal drugs in one product. There is potential risk from impurities/adulterations of herbal medicine mixed with toxic plant extracts because of misidentification or intentional addition of allopathic drugs.

The European pharmacopoeia includes 120 monographs on herbal drugs. Control of the starting materials is essential to ensure the reproducible quality of herbal medicinal products. Herbal drugs must be accurately identified by macroscopic and microscopic comparison with authentic material. Herbal drugs are referred to by their binomial Latin names of genus and species; only permitted synonyms should be used. Different batches of the same herbal ingredient can differ in quality because of a number of factors.

- Inter- or intra-species variation.
- Environmental factors.
- Time of harvesting.
- Plant part used—active constituents usually vary between plant parts, and it is not uncommon for a herbal drug to be adulterated with parts of the plant that are not normally used.
- Storage conditions and processing treatments can greatly affect the quality of an herbal ingredient.
- Instances of herbal remedies adulterated with other plant material and conventional medicines.
- Extraction/drying methods.

Identity tests establish the botanical identity of a herbal drug.

- Chemical (e.g. colour or precipitation) and chromatographic tests are used for identification of the ingredients.
- Assay—a herbal drug with known active principles should have an assay established to set the criterion for the minimum acceptable percentage of active substance(s).

Legislation of herbal drugs

Although herbal drugs have been used as traditional remedies for centuries and are perceived by many to be without major safety problems, the UK has a series of controls to limit general availability.

Hazardous plants, such as digitalis, rauwolfia, and nux vomica, are specifically controlled under the Medicines Act as prescription-only medicines (POMs).

Certain herbal ingredients are controlled under the Medicines (Retail Sale and Supply of Herbal Remedies) Order, 1977, SI 2130. This Order (part I) specifies 25 plants that cannot be supplied except by a pharmacy,

and includes well-known toxic species such as areca, crotalaria, dryopteris and strophanthus.

Herbal remedies exempt from licensing fall under two main categories:

- Subject to the provisions of section 12 of the Medicines Act 1968, products can be compounded and supplied by a herbalist on their own recommendation.
- If no medical claims are made that are attributable to the herbal product, it can be sold as a food supplement.

Efficacy

Herbs used medicinally normally have a traditional reputation for their uses, but generally there is little scientific documentation of their active constituents, pharmacological actions, or clinical efficacy.

The current emphasis on EBM requires evidence of efficacy from rigorous randomized controlled trials. Several systematic reviews have been prepared by the Cochrane Collaboration. These reviews highlight that, in some cases, the evidence base is weak and studies are often flawed. Evidence from randomized controlled trials has confirmed the efficacy of St John's wort products versus placebo in the treatment of mild to moderate depression.

If the active constituents of a herbal drug are known, it is possible and, in most cases, desirable to standardize the extract. The aim of standardization is to obtain an optimum and consistent quality of a herbal drug preparation by adjusting it to give a defined content of a constituent or group of constituents with known therapeutic activity. Examples include senna, frangula, digitalis, belladonna, and horse chestnut.

In the case of St John's wort, early studies concentrated on the hypericin constituents, but more recent work suggests that hyperforin and, possibly, flavonoids also contribute to the antidepressant properties.

Safety and adverse effects

Information on herbal medicines is lacking in many areas including active constituents, metabolites, pharmacokinetics, pharmacology, toxicology, adverse effects, long-term effects, use by specific patient groups, and contraindications.

- Herbal drugs could present a potential risk to health from exposure to contaminants present in the herbal product and result in ADRs (Table 8.1).
- Reliance on self-administration of herbal drugs or products could delay a patient seeking qualified advice or cause a patient to abandon conventional treatment without appropriate advice.
- In some cases, herbal medicines could compromise the efficacy of conventional medicines through herb–drug interactions.

Table 8.1 Adverse reactions associated with herbal medications

Adverse reaction	Herbal medication
Cardiotoxicity	Aconite root tuber, ginger, licorice root, mahuang
Cross-sensitivity with ragweed	Arnica, calendula, dandelion, echinacea, feverfew, german chamomile, golden rod, march blazing star, milk thistle, mugwort, pyrethrum, stevia, tansy, wormwood oil, yarrow
Gastrointestinal (nausea, emesis, dyspepsia, etc.)	Echinacea, ephedra, evening primose oil, garlic, ginger, milk thistle, soy
Hepatotoxicity	Borage, calamus, chaparral, Chinese herbs, coltsfoot, echinacea, germander, kava rhizome, kombucha, life root, mahuang, pennyroyal, sassafras, skullcap, soy, valerian
Neurotoxicity	Aconite root tuber, ginkgo seed or leaf, kava rhizome, mahuang, penny royal
Renal toxicity	Chinese yew, hawthorn, impila root, penny royal, star fruit
Sedation	Chamomile, ginger, St John's wort, valerian

This page intentionally left blank

General information about commonly used herbal medications (Table 8.2)

Table 8.2 General information about commonly used herbal medications

Herbal drug	Use(s)	Proposed mechanism of action	Other considerations
Black cohosh	• Treat PMS and dysmenorrhea • Reduce menopausal symptoms such as hot flushes	• Possibly has oestrogen-like activity • Suppresses lutenizing hormone secretion	• Not recommended for >6 months • May relieve vasomotor symptoms • Effect on breast cancer, osteoporosis and cardiovascular risk is not known
Chamomile	• Reduce anxiety and insomnia • Relieve GI spasms	• Contains flavonoids which are the active component • Benzo-diazepine receptor binding ligand	• Allergic reactions reported esp. if patient has ragweed allergy • Sedation is additive with other therapies
Echinacea	• Treat and prevent colds • Stimulate the immune system	• Increases phagocytosis and lymphocyte activity • Anti-inflammatory	• Not recommended for longer than 8wks • Can worsen asthma
Evening primrose oil	• Treat symptoms of PMS and menopause	• Active component probably linoleic acid	• Evidence is controversial • Side effects include headache, nausea, and diarrhoea
Feverfew	• Prevent migraines • Relieve dysmenorrhea • Improve inflammatory processes	• Inhibits prosta-glandin synthesis • Analgesic properties	• Rapid with 'post feverfew syndrome' which includes anxiety, headaches, and insomnia • Must be taken daily for migraine prevention • Not used for migraine treatment • Contraindicated in pregnancy

Table 8.2 *(Contd.)*

Herbal drug	Use(s)	Proposed mechanism of action	Other considerations
Garlic	• Lower cholesterol • Treat hypertension • Prevent stomach and colon cancer	• Antioxidant and antiplatelet activity • Smooth muscle relaxant and vasodilator • HMG-CoA reductase inhibitor	• Odourless preparations have less of the active component • Enteric coating ensures proper absorption
Ginger	• Decrease GI upset and nausea • Reduce post-surgical nausea	• Serotonin antagonist at 5-HT3 receptor in ileum • Anti-inflammatory	• Toxicity includes sedation and arrhythmias • Adverse effects include gas, heartburn, and bloating
Ginkgo biloba	• Enhance memory • Treat or prevent dementia	• Antioxidant • Increases blood circulation by decreasing viscosity • Regulates vascular smooth muscle	• Uncooked seeds contain ginkgo toxin which can cause seizures
Ginseng	• Stimulate the immune system • Improve blood glucose and BP control	• Increases cortisol concentrations • Stimulates natural killer cells	• Limit use to 3 months • May cause sleep disturbances • Avoid large amounts of caffeine
Hawthorn	• Treat heart failure • Improve hypertension • Treat coronary heart disease	• Anti-inflammatory properties • Lipid-lowering properties	• May decrease dyspnoea and fatigue • No mortality or morbidity data
Horse chestnut	• Improve symptoms of chronic venous insufficiency • Decrease leg oedema	• Seeds contain aescin which reduces venous capillary permeability • Anti-inflammatory • Weak diuretic activity	• May increase bleeding when in combination with warfarin • Can turn urine red • Can cause kidney or liver damage
Licorice	• Treat stomach ulcers • Relieve constipation	• Glycyrrhizin and glycyrrhetinic acid prevent the degradation of prostaglandins in the gastric mucosa • Antioxidant activity	• Can cause sodium and water retention and hypokalemia • Avoid in patients with cardiovascular or renal disorders

(continued)

Table 8.2 (*Contd.*)

Herbal drug	Use(s)	Proposed mechanism of action	Other considerations
Milk thistle	• Protect the liver	• Seeds contain silymarin • Antioxidant, anti-inflammatory activity • Inhibits mitochondrial damage	• GI side effects are common including nausea, diarrhoea, and fullness • Cross-sensitivity to ragweed allergy
Pepper-mint	• Reduce nausea and indigestion • Treat headaches • Improve irritable bowel syndrome	• Direct relaxing on GI smooth muscle • Inhibits potassium depolarization in intestine	• Avoid in patients with pre-existing GI disorders • May decrease absorption of iron
Saw palmetto	• Treat benign prostatic hyperplasia	• May inhibit 5-α-reductase • Local anti-androgenic and anti-inflammatory effects on prostate	• Symptom improvement similar to that seen with finasteride • No long-term data
Soy	• Decrease cholesterol • Relieve menopausal symptoms • Improve bone mineral density	• Isoflavones bind to α and β oestrogen receptors	• Causes nausea, bloating, and constipation
St John's wort	• Treat depression and anxiety	• Active components, hypericin and hyperforin, inhibit serotonin, dopamine, and norepinephrine re-uptake	• May cause photo-sensitivity • Avoid in patients with psychiatric illness including bipolar and schizophrenia • May have withdrawal effect after chronic use
Valerian	• Treat anxiety and insomnia	• Binds with GABA receptor in CNS	• Can cause excitability with high doses • Takes weeks for effect

Chinese herbal medicine

Most of the substances used in Chinese herbal medicine originate from China. The Chinese pharmacopeia lists over 6000 different medicinal substances; there are currently over 600 different herbs in common use. Herbs are used for their abilities to treat specific Chinese diagnoses and alleviate specific complaints. For example, there are assortments of herbs that can alleviate coughing, but each one is appropriate for a cough with a different Chinese diagnosis. The variety and degree of different combinations of herbal medicines makes Chinese herbal medicine very complex.

Combination of herbal products

The one characteristic of Chinese herbal medicine that most differentiates it from other types of herbal medicine is the degree of combination undertaken. Chinese herbalists very rarely prescribe a single herb to treat a condition; instead, a mixture could contain >20 herbs. Pre-prepared formulas are available; however, these products are not usually as potent as the traditional preparation of 'decoction'.

Decoction is the traditional method of preparing herbal medicine. A decoction is a concentrated form of tea. The practitioner weighs out a day's dosage of each herb and combines them in a bag. A patient is given a bag for each day the herbal formula must be taken. The herbs are then boiled in water by the patient at home; the boiling process takes 30–60min and the resulting decoction is consumed several times during the day.

Quality issues

The quality and safety of Chinese herbs has repeatedly come into question after media coverage of concerns over heavy-metal contamination, adulteration, and use of endangered animal species. Heavy-metal contamination has been detected in several Chinese herbal products, usually as a result of poor manufacturing. Adulteration of herbal medicines with prescription drugs has been found in a few herbal products. The use of endangered animals in Chinese herbal medicine is very rare.

Herbal interactions

Information on herb–drug interactions (Table 8.3) is generally limited to case reports, although recognition is improving, with the result that clinically important interactions are increasingly being identified and prevented by healthcare professionals.

Variability of constituent ingredients and the pharmaceutical quality of unlicensed herbal products can often be the main reason for the low incidence of reported interactions.

Types of interaction

Pharmacokinetic interactions with drugs

- Absorption
- Distribution
- Metabolism
- Excretion

Pharmacodynamic interactions with drugs

- One substance affecting the response of another at its site of action.

Herb–disease interaction

Certain underlying diseases could be exacerbated by ingestion of herbal ingredients with the following properties:

- hypertensive properties.
- hyperglycaemic/hypoglycaemic activity.

Table 8.3 Some important herb–drug interactions

Herb	Drug interaction	Considerations
Black cohosh (*Actaea racemosa*)	Antihypertensives	May ↓ BP
Chamomile (*Chamaemelum nobile*)	Anticoagulants	Consider discontinuing 2wks before surgery
Echinacea purpurea	Immunosuppressants (eg corticosteroids)	Immune suppression can result from prolonged use for >14 days Loss or ↓ in therapeutic effect of some drug therapies; probably induction of CYP enzymes
Ephedra (ma huang)—active constituent is ephedrine	Will have the same interactions as ephedrine	Misuse has resulted in death
Evening primrose oil	Could interact with anti-coagulants or antiplatelet drugs Can ↓ seizure threshold	

Table 8.3 (*Contd.*)

Herb	Drug interaction	Considerations
Feverfew (*Tanacetum parthenium*)	Could interact with anti-coagulants or antiplatelet drugs	Consider discontinuing 2wks before surgery
Fish oil supplements (omega-3 fatty acids)	Reports of ↓ platelet aggregation	Unlikely to have clinical significance
Garlic (*Allium sativum*)	Could interact with anti-coagulants or antiplatelet drugs	Consider discontinuing 2wks before surgery Loss or ↓ in therapeutic effect of the drug therapies; probably induction of CYP enzymes
Ginger (*Zingiber officinale*)	Could interact with anti-coagulants or antiplatelet drugs	Consider discontinuing 2wks before surgery
Ginseng (*Panax ginseng*)	Could interact with anti-coagulants or antiplatelet drugs Interacts with hypoglycaemic drugs Avoid concurrent monoamine oxidase inhibitors (MAOIs)	Varying effects on BP Hypoglycaemia Could potentiate action of MAOIs Limit use to 3 months
Hops (*Humulus lupulus*)	Could have additive effect with CNS depressants	Avoid in depressive states
Horse chestnut (*Aesculus hippocastanum*)	Could interact with anti-coagulants or antiplatelet drugs	↑ risk of bleeding
Passion flower (*Passiflora incarnate*)	Additive effects with CNS depressants Avoid concurrent MAOIs	Reports of hepatic and pancreatic toxicity
Saw palmetto (*Serenoa serrulata*)	Caution with finasteride	Potential of additive effect
St John's wort (*Hypericum perforatum*)	Anticonvulsants Ciclosporin Digoxin Protease inhibitors and non-nucleoside reverse transcriptase inhibitors (NNRTIs) Oral contraceptives Theophylline Warfarin Irinotecan	Loss or ↓ in therapeutic effect of the drug therapies; probably induction of CYP enzymes by St John's wort constituents
Valerian (*Valeriana officinalis*)	Additive effects with CNS depressants	
Milk thistle (*Silybum marianus*)	CYP3A4 enzyme inducer Protease inhibitors and NNRTIs Phenytoin	↓ blood levels and hence chance of treatment failure

Please note that this is not an exhaustive list but a point of general reference. New information about herbal interactions can be obtained from http://www.mhra.gov.uk

Perioperative considerations for herbal drugs

- Herbal medicines have the potential to pose problems in the peri-operative setting because patients often fail to communicate concurrent herbal remedies during DHx taking by healthcare professionals.
- Few data exist in the medical literature regarding the use of herbal products and the development of ADRs or interactions associated with anaesthesia.
- The most important risks associated with herbal products during the perioperative and immediate postoperative periods are cardiovascular, coagulation, and sedative effects.
  - Cardiovascular effects—ephedra, ginseng, and garlic:
    - — ephedra can cause a dose-dependent ↑ in heart rate and BP.
    - — ginseng ↑ BP and its use is not recommended during the surgical period in patients with cardiovascular disease.
    - — garlic could ↓ BP, but its effects are normally brief and usually require high dosages.
  - Bleeding effects—garlic, ginseng, gingko, evening primrose oil, feverfew, fish oils, ginger, horse chestnut, and kava kava.
  - Sedative effects—chamomile, kava kava, valerian, hops, passion flower, and St John's wort.

Although there continues to be debate on the incidence of reactions to herbal products during the perioperative period, it might be prudent to recommend discontinuation of these agents for at least 2wks before surgery.