How to write a scientific paper

©Ronán M Conroy
Royal College of Surgeons

Introduction

A scientific paper has a form as abstract and rigid as a sonnet. This may seem guaranteed to cramp your style, but it actually makes life easier for scientists. No matter what your speciality, you will be able to find your bearings in any scientific paper. This is important. With the medical literature, for example, expanding at a rate equivalent to three quarters of a kilometre of bookshelves per year, scientists have to be able to evaluate the relevance, significance and quality of a paper in as little time as possible.

The reason I decided to write this paper was because no-one ever explained to me how to go about writing a paper. I knew, of course, that I needed an introduction, a method section, a results section and a discussion, but after that I was on my own. More years and more scientific papers than I care to think about later, I wish someone had given me a little more guidance. My work often involves helping people to write their first papers. The guidelines in this paper are the result.

The scientific paper: gross anatomy

A scientific paper has four parts:

- **The introduction**
  - This says *why* you did the study
- **The method section**
  - This says *how* you did the study
- **The results section**
  - This describes your findings
- **The discussion**
  - This places your findings in the context of the evolving research literature.

This may seem self evident, but the first problem you will encounter is how to write your paper maintaining these distinctions. The introduction can easily degenerate into a free-for-all discussion of the current literature, and the results section may provoke a fierce desire to begin commenting on the significance of the results. It seemed odd to me, at first, that I could not explain the more difficult parts of the statistical methodology in the context of the actual data analysis, where they would be easier to understand. I now regard writing a good general description of the statistics methods as a challenge, knowing how much I have learned from reading such descriptions over the years.

Although the abstract is the first thing to appear, you should write it last. It summarises the paper, and to write it you will need to be able to refer to the text of the paper. I will come to it last.

The Introduction

The ideal introduction has three sentences:

- This is what we already knew.
- This is what we didn't know.
- To fill this gap, we decided to do this study.

The introduction situates the research in the context of what is known and unknown in the field. It should be brief—three sentences is ideal, anything more is a compromise. Its function is to act as a shop window. The reader must be able to decide whether your report is interesting and relevant based on the introduction. Do not tax their patience.

Let me say something about each of the three 'sentences'.

1. **Stating what we already know**

   Assume that your reader knows the area. (You will get a chance, in the discussion, to give a reasoned overview.) For the moment, marshal the salient findings as briefly and simply as possible. For example

   > *Studies of frequently-attending patients have shown that they have significantly higher levels of health anxiety than other patients (references).*

   There is no written rule, but when referencing the introduction, I would restrict myself to at most three of the most relevant and recent articles. At the end of the sentence above I would have put one, or at most two references to recent review articles. I would have put two references only if the phenomenon had been studied from two different perspectives and, as a result, had developed two divergent literatures (this often happens in areas in the no-man's land between disciplines, such as psychology and health).

   Having summarised (rather brutally!) the state of current knowledge, the next sentence is

2. **Stating what we do not know**.

   This should follow on logically. To continue the example,

   > *It is not known whether levels of health anxiety are related to frequency of attendance among the majority of patients: those who are neither frequent attenders nor suffer from abnormally high levels of health anxiety.*

   This justifies doing the research. We know about the pathological case, but we wonder if the same mechanism hold true within the 'normal range'. Having made a case for doing the study, we then describe the study we did.

3. **Stating the conceptual question**

   Before a study can be carried out, the conceptual question that it asks must be translated into **operational** terms. A conceptual question is simply the question in plain everyday language; the operational question translates the vague terms in the question into measurable terms. To do research, we must decide how we will measure things, how we will select participants, how we will exclude bias and so on. However,

all of these decisions need to be guided by an overriding principle: the study must be the best possible test of the conceptual question. In order to decide whether the study is a good test of the conceptual question, we need to write down the conceptual question. (And I would urge you to write it down when you are planning the study, and not to have to puzzle it out by reverse engineering afterwards!).

The conceptual question, then, puts the study question in plain language. It puts aside, for the moment, the problems of how to measure, how to sample, and so on. For instance

> *We assessed the relation between health anxiety and health service use in a series of general practice patients.*

We have not defined how we are going to measure health anxiety, nor said which characteristics we are going to measure (though frequency is a fair bet!). All this will be done in the method section.

I know: I have chosen an example where the whole introduction could be written in three sentences. For some papers, each sentence may end up being a paragraph. But the principle is the same. I advocate writing as tersely as possible and adding detail only under duress in the

The conceptual question forms the last sentence of the introduction of any scientific paper. Get hold of papers and read the conceptual questions as an aid to helping you to write yours. The conceptual question is the single most important sentence in the paper. The whole logic of your paper will flow from it. Without a clear statement of the conceptual question, the reader cannot decide if you were studying the right people, asking the right questions, analysing your data right or wrong — nothing. So it's the thing that needs most polish.

The method section

The method section describes all the decisions that you took in order to make the conceptual question a research question. You had to take decisions about who to study, what information to collect and how to present it. The method section details how you translated your question into practice. There are two reasons why you need to describe this process:

1. Someone may want to **replicate** your study. They will need to be able to do so from your description.
2. A much more important reason, however, had to do with **relevance**. Any study is only one of many possible studies that could be done to test the conceptual question. In order to judge the contribution of your study to illuminating the conceptual question, the reader must know how you defined your terms. This is easy enough when you are describing cardiovascular mortality. In this case, you simply need to specify how deaths were identified and how cardiovascular were separated from non-cardiovascular. Things get more complicated when you describe, say, cardiovascular morbidity. Is an angioplasty a non-fatal cardiac event or not? Depending on your

decision, your event rates will change, and the predictor factors may gain or lose strength of association. It is vital to the interpretation of your study that you specify exactly what constituted a cardiovascular event for the purposes of the experiment. How much harder still is it to define what we mean by health anxiety!

These two criteria — replicability and relevance — guide the writing of the method section. How much detail do you need to go into? This is hard to say exactly (and you will inevitably get a request from at least one referee for more detail and from another for less!). Here are a few guidelines:

- Mention well-known procedures by name. You do not need to describe or reference the t-test, or how Body Mass Index is calculated. These are familiar to most readers.
- Mention specific equipment or methods only when there are several competing alternatives, the choice of which would affect your result. You should mention what ambulatory blood pressure monitor you used, for instance, since there are differences between various types.
- When mentioning a procedure which is not familiar, give some detail on how well it works and where the interested reader can learn more. In the case of health anxiety, for instance, I would mention who developed the scale, what studies had been done to validate it, and whether it had fared well in reliability studies.
- Do not mention statistical procedures that are commonplace. (Some people mention the t-test and the chi-squared test in their method section, but don't mention that the blood sample was taken with a needle. My feeling is that if there is one obvious way of doing it, you only need to specify exceptions.)
- When you mention a test, procedure or method that is not widely-used or which you are using in a novel way, give a reference to a good review article, ideally one which discusses the application to your area of study. This is important. If someone reads your paper and becomes excited at the possibilities of a methodology which you have used, the reference should send them to an article that they can understand. Therefore, choose one that is, if possible, geared to the sort of readers you expect to read your paper.
- Do not content yourself with a good reference, however. If you introduce a novel method, provide a short description that will allow the reader to make sense of your paper. I might write something like:

  > *Multinomial logistic regression is an extension of logistic regression. Logistic regression is used to examine the factors which predict of a single endpoint which either occurs or does not. Multinomial logistic regression extends this to situations in which one of a number of different endpoints can occur, but only one. In the present study, patients could either develop gestational hypertension or pre-eclampsia, or remain normotensive. The latter group constitute the baseline group against which the relative risk ratios are measured.*

- If you use a scale which was developed by someone else, but which would not be familiar to a lot of readers, then it is useful to say what it measures, how reliable it is,

and to give a sample item or two, to give the reader a 'flavour' of the scale.

- If you design your own questionnaire or data collection instrument, include it in the paper as an appendix. It may not be published (in fact, it is unlikely to be, in my experience) but it will greatly help the journal reviewers to appreciate where your data came from and how it was collected. When submitting a paper, I always point out that I have included the questionnaire as an appendix to help the reviewers, and (diplomatically) add that I will be guided by their advice as to whether it should be published with the paper or not. My experience as both an author and a reviewer is that this goes down well.
- Finally, I often give details of the statistics package used, as it is handy for researchers to know what packages can do what.

In order to judge the relevance of your report, there are two particular issues that a reader will have in mind:

- Were the measures used reliable and valid, and
- Could bias have entered the study at any point?

With that in mind, you should anticipate questions such as how you ensured that the sample really was representative, or how you ensured that all follow-up events were detected, or how you prevented bias from entering the selection of people for the study or the measurement of the variables.

The results section

When I say that the results section should give all the results and only the results, you will probably feel I am labouring the obvious, but this is one of the things that I and many other people still find hard about scientific writing. There is an urge to comment or discuss the significance of the findings right away. It is best avoided. The idea of separating your results from your discussion is that the reader should be allowed to inspect your results in peace before being conducted on a guided tour by you, complete with your explanations and interpretations, in the discussion section.

Writing the results section requires a lot of organization. The thing never ever to do is to amass a whole heap of computer output and then put an evening aside to organise it. This is just like descending on the shops and buying anything that you see that looks good to eat and then going home and trying to figure out what to cook. Without a ground plan of the results section, you cannot decide on what are relevant and irrelevant results.

Like choosing your recipe before you go shopping, I make researchers hand me a results section before the data analysis. The first time someone asks you to do this, it's rather disconcerting. How can you write a results section without results? Actually, it's the only way.

Without the encumbrance of tables of 'everything broken down by age and sex' (a request that every statistician knows to ignore) you can think of a logical order for the results. While there is no universal template you can follow for a results section, many studies follow what I call the three table plan.

The three table plan

In this plan, you think of your results as three tables. Table 1 describes who was studied; table 2 describes what happened and table 3 examines the relationship between who you are and what happens to you.

Table 1: who was studied

Table 1 gives the characteristics of the people studied. (I am using human studies as a template—your study could equally well be about rats, accidents, sphygmomanometers or earthquakes). The purpose of table 1 is twofold:

Relevance: the person reading your paper should be able to decide whether the people you studied are relevant to them. The decision to study a particular population is one of the fundamental decisions that you make when moving from a conceptual question to an operational one. And it is a critical one. Remember that, for many years, hypertension in older people was treated with drugs that were at least less than satisfactory and at worst dangerous, all because treatment studies of hypertension had been confined to middle-aged people. Your paper describes the investigation of the research question in the context of a specific population. If this population is of limited relevance, table 1 will show it.

Bias Bias is any feature of the design of your study which can cause the results to be systematic under- or over-estimations of their true values. In a scientific experiment, many of the features of the design of the study are geared towards eliminating systematic differences between groups that are being compared. For instance, in a study of a new treatment, patients are allocated to the new or old treatments at random, thus making sure that there is no systematic difference between each treatment group. In an observational study, however, we cannot allocate people at random. So it is that, for instance, if you study women who smoke during pregnancy, these are also the women who are likely to drink and to use recreational drugs; they are also more likely to be unemployed. So a comparison of the blood pressure of the women who smoked with the women who did not may produce a biased estimate of the effect of smoking, as it also includes the effect of a higher alcohol consumption and the effect of unemployment (both of which are associated with blood pressure levels).

Table 1 should serve two purposes, then. It should give enough information about the people you studied to allow a reader to judge how relevant you study was to their interests, and it should report on the factors which could bias the conclusions of the study.

The information in Table 1 needs to be organised in a logical order of some sort. I favour a chronological one. I start with the basic characteristics like age, sex and family history, then go on to later ones like age at completion of education, age at onset of disease or previous medical history. The final group of items is the one that describes the person at the time of the study, such as presenting problem, current disease status an so on. Another strategy is the macro-to-micro approach where you start with the most global characteristics and go on to the more detailed ones.

Table 2: what happened

Table 2 shows the data that relates to the research question that prompted the study. If your study is purely descriptive (such as a survey of opinions) then table 2 shows the breakdown of answers to each of the items surveyed. If your study is examining a relationship, then the table shows it. For instance, if you are studying blood pressure and risk of complications of pregnancy, then table 2 shows each complication in relation to blood pressure levels. If you are reporting a treatment trial, table 2 shows the trial endpoints overall and within each treatment group.

Table 3: the relationship between who was studied and what happened

In observational (non-experiemental) studies, tables 1 and 2 naturally provoke the question 'could the results in table 2 be biased by the characteristics in table 1?' Even in treatment trials, we could well ask 'did the treatment work equally well in all patient subgroups?'. Table 3 attempts to answer these questions. It often presents the results of procedures with 'multiple' in their names — multiple regression for instance. The question that is frequently addressed in table 3 is 'what is the estimate of the results shown in table 2 when we correct them for the biasing effects of factors in table 1?'

For instance, women who continue working in pregnancy have higher blood pressure levels than women who stop or women who are unemployed. They also have a higher risk of pre-eclampsia. They also differ in their smoking, alcohol consumption, age and even in their marital status. Could these differences be responsible for the apparent difference in blood pressure. And because pre-eclampsia is characterised by high blood pressure, we could also ask if the blood pressure differences between the working and non-working women were enough to explain the difference in risk of pre-eclampsia. This is table 3 stuff. I would tackle the last question, for instance, by calculating risks of pre-eclampsia unadjusted, then adjusted for age, smoking, alcohol consumption and marital status, and finally also adjusted for blood pressure.

The three table approach doesn't necessarily mean that you end up with three tables! For one thing, there may be no reason to have a table 3. There may be no interesting questions to ask about how the characteristics of the people studied would have affected the estimate of the results in table 2. But in any event, when I talk about three tables, this is a very useful way of organising the results to give yourself a good overview, but it may not be the optimum way of publishing them.

Writing the results section

Once you have your results as tables, you can write the results section. While you are doing so, you can be considering other ways of presenting your data, but more of that later. The first problem is to walk the reader through your results, which is what a results section should do. Your writing should be a continuous narrative as you conduct the reader through your findings. They may break off to look at a table, but you should think of the prose in the results section like a good guided tour.

Avoid the temptation to read the table aloud to the reader. *There were 34 women*

(18.2%) with a previous history of hypertension, who had a mean age of… Try to be more helpful. Ask yourself what you would like the reader to notice in the table, and try and present that information using as few numbers as possible. I would write something like *almost a fifth of the patients had a previous history of hypertension; these patients were, on average, 5.1 years older than the patients with no previous history*. The interested reader can refer to the table to find out the actual mean ages of the groups, or the exact percentage, while the salient information — an age difference of potential significance for the results— is all that the reader must be aware of at this point.

As you write your results section, you may find that some of the figures presented in the table can be drawn effortlessly into the writing. In a number of papers, table 1 has disappeared altogether and been replaced with a short paragraph of writing. Don't feel obliged to keep the table just because you prepared it! As an exercise in style, I once wrote a paper which had no tables or graphs. I felt that in the particular case the results were better presented as a continuous narrative, and the journal editor agreed with me. So don't try to make your results section look like everyone else's — try to make it the best presentation of your results.

As you write, it also may become clear that a table is better split into several smaller ones, or that a table can be simplified by writing some of the results in the text. One case in point is where there are no differences between groups. In this case, present only the overall results and mention in the text that there was no difference.

The other decision to be made as you write is whether a graphic might be a better way of presenting the data.

Tables, text or graphics?

Tables present results numerically. This has the advantage of the precision that numbers bring, but it also requires an effort from the reader. Numbers are symbols. They do not look like what they represent. There are two alternatives, each with its own advantages. Words are a useful, and sometimes overlooked, way of presenting results. Words allow you to keep up a continuous narrative, which helps to present the results in the best order and to highlight their significance. Graphs have the advantages that they are processed by visual perception, unconsciously. A person may fail to notice that the standard deviations of two groups differ remarkably — most readers ignore standard deviations in tables, usually because they cannot interpret them, so I don't usually report them, on the grounds that if there is something important about the scatter of the data, I should think of a clearer way of presenting it. In this case, I would present two boxplots, which not only show us whether the scatter is similar in two sets of data, but also tell us why standard deviations might be different (a larger standard deviation can result from skew, from outliers or from a distribution of similar shape but wider scatter; the standard deviation does not help to decide which explanation holds).

Since graphics are processed unconsciously, the reader cannot help but see the pattern in a good graphic, whereas they have to decipher the pattern for themselves in a table. This is not the place to go into the choice of graphics, except to say that statistical graphics has flowered as a discipline in the past two decades, and that researchers who confine themselves to pie charts and histograms are doing the equivalent of trying to practice diagnosis with a

stethoscope and a thermometer.

The discussion

The discussion is a four part affair. It should cover

1. what we knew before this study,
2. what we found in this study,
3. what we now know and do not know (scientific implications), and
4. what we should do or should not do (practical implications).

The first section is easy enough. Now is the time to write a masterly review of all those articles you amassed as a result of your expertise searching the literature. Be careful, however; you are not writing a review paper. Limit yourself to the literature directly relevant to your research question. I admit that if I find a good review article on the wider issues, I mention it and give a reference. I do so because I would be glad of the reference myself. But be disciplined; if you are writing about use of complementary and alternative medicine in cancer patients, your focus is cancer patients. Beyond that, there may be useful leads from research into patients with other life-threatening diseases, which can help to situate the research (do cancer patients resemble other types of patients studied). But any extension of the discussion to things not studied in the paper needs justification.

The second section is somewhat trickier. You have already presented the results; now is your chance to say what they mean. Now you are free to interpret patterns in the data. Now that your have given your reader a guided tour of the data, you can interpret them and comment on their scientific significance. Do not quote numbers in the discussion unless they are striking and need to be recalled to mind. You can say things like *there continued to be a strong association even after differences in age and initial blood pressure were controlled for statistically*.

Finally, you must integrate your findings into the accumulated research literature. In what way do your findings confirm previous work? Do they extend previous work into any area not previously studied, such as different groups of people or settings? Do they disagree with any of the previously published reports? If so, could that be due to differences between the way you ran your study and the previous people ran theirs? (Differences in who was studied, how they were studied, methods used to analyse the data?).

And now, you get your chance to take the podium! What are the practical and theoretical implications of the research so far, and what are the questions that need to be answered next? This is the only area of a paper in which you can speculate, so enjoy it. But don't forget it, either. Many papers end with a dull summing-up of the research literature. Surely someone who has been working in the area, thinking about it, reading about it, has some ideas! Don't sell yourself short. You probably know more about this topic than the great majority of your readers (put modesty aside for a moment; this is likely to be true). For that reason, you should have some interesting opinions. Express them. And if you don't have any, then maybe you should think some more before finishing the paper.

The Abstract

The abstract is becoming the single most important part of a paper. This is because many electronic literature search systems can display the abstract of a paper. Abstracts were once qualitative summaries of research papers. Now they are mini-papers in themselves. You should read the instructions for authors of the journal you are aiming for, because many journals now require structured abstracts. A structured abstract has a short paragraph corresponding to each section of the paper: introduction, methods, results and discussion. In fact, you should write a structured abstract anyway. If the journal doesn't use headings in the abstract, you can delete them.

Since the abstract is a mini-paper, you should follow the structure for each section closely. However long your **introduction**, the corresponding paragraph in the abstract should be no more than three sentences, following the sequence I described above. The **methods** section in the abstract should say

- what sort of study this was
- who was studied
- what were the main measures

The **results** section of the abstract should concentrate on results of importance. You can skip most of the 'table 1' results, since they just describe the group you studied. Give the bare minimum here in order to allocate as much space as possible to the main findings. Try to summarise the main findings without cluttering them with p-values and forests of numbers. Simply report findings that were in line with previous research. For instance, you can say

> *Health anxiety scores declined modestly with age, and were similar in men and women.*

Select the results that you most want to highlight and present them numerically. For instance

> *Frequency of visiting rose from 2.6 per year in the lowest decile of health anxiety to 4.2 in the highest (p = 0.033).*

This highlights the finding (which was a first) and tries to express it in as practical terms as possible. (Note that a lot of the impact would have been lost if I had just said 'there was a significant relationship, and given a p-value.)

Finally, the **discussion** section of the abstract should focus on what the study contributes to the literature. For instance (staying with the same paper)

> *In this population, absolute levels of health anxiety were low. Nevertheless, within this 'normal' range, there was an association between level of health anxiety and frequency of visiting, and with expectations for the index visit. The findings support the health anxiety model as part of the dynamics of services use by psychologically normal persons.*

If someone orders a reprint of your paper, it will probably be on the basis of the abstract. And, indeed, the abstract will probably be the first thing a journal reviewer reads. Write it carefully. I often write my abstract by copying and pasting sentences from the paper itself and then editing and shortening them. This guarantees that the abstract agrees with the text and, where possible, uses the same wording. For the same reason, I often copy text directly from the abstracts of papers that I am citing rather than paraphrasing the paper myself. This makes sure that the authors are reported in their own words.

STYLE

It only remains to say a word about style. Scientific papers are often written in a very roundabout, impersonal style. It is important to remember that although English has become the language of science, many readers will not speak it as a first language. You should avoid needlessly complex language. Do not be afraid to use words which seem to belong to conversational English rather than English prose. A technical editor I admired greatly once replaced the word 'therefore' with the word 'so' in a paper I published in the British Heart Journal, and after my initial shock I was delighted with the result.

Avoid impersonal constructions. These used to be the norm in scientific papers, but they can lead to really awkward writing. Saying *All patients presenting with chest pain who had a previous history of gall-bladder investigation were studied* is vague. Who studied them? Why — you did! Then write *We studied all patients who presented with chest pain and who had a previous history of….*

Avoid acronyms. Many TLAs (three-letter acronyms) that are used in papers simply serve to make the text harder to read. I restrict myself to using only the acronyms that have become general and require no explanation, such as IQ and GP. Each discipline has its own acronyms which can be daunting to the non-specialist. A psychiatrist, for example, will recognise PTSD, and often use the term in what passes for conversation, but may well be perplexed when an obstetric colleague refers to CPD. Better that they both write about post-traumatic stress disorder and cephalo-pelvic disproportion and have done with it.

Some people justify acronyms on the grounds that they save space. I doubt if the world's forests will be greatly helped by using acronyms, and the shortening of the text is usually at the cost of making it much harder to read. Several journals do not allow acronyms, and after my first experience of working without them I have been their tireless enemy.

Finally, some curious customs. When referring to things that are accepted as true on the basis of the findings of others, use the present tense. For example, *smokers have a lower in-hospital mortality after an acute coronary event than non-smokers (reference)*. The only time you use the past tense is when you refer to the findings themselves, such as *Barrett-Connor and her colleagues found a 10-year mortality of 42% (reference)*. I use the word colleagues if I happen to know the sex of the main author, just to vary the tiresome *et al*. When you refer to your own findings, you use the past tense. *Women who continued to work had higher daytime pressures than those who stopped*. This habit is a sort of British Modesty that goes with the use of English. The present tense is only used for describing things believed to be true, and you never use it to describe your own findings, since it is up to your colleagues to accord them that status. Well, that's what I was told.

My mentor, Risteard Mulcahy, ruthlessly removed all words like 'interesting' from my drafts, saying, with that wry smile of his 'sure, everything we publish is interesting.' I am not so sure, but I still use words that reflect value judgements very sparingly. (I am probably over-fond of the word 'curiously'.) I often find myself pruning them from other people's drafts. I think that if a paper tells its story well, there is no need to draw attention explicitly to what is remarkable.

Making the best use of your computer

If you are writing a paper, you should get to know the capabilities of your word processor. One important capability is **outlining**. This allows you to structure a document with headings and subheadings. If you use a word processor and cannot use outlining, then learn. It is a valuable aid to coherent report-writing.

You can add references to your document as footnotes. This has the advantage that the references are numbered automatically, and moving or deleting a reference will automatically cause renumbering. At a more sophisticated level, bibliographic packages such as Endnote, ProCite and Reference Manager offer dedicated databases for keeping your references, including the ability to create a database direct from an electronic literature search. These packages will add references, correctly formatted, from the database to your paper. Invaluable for the serious researcher, though expensive and slow to learn if you only write the occasional paper. ProCite, though still available, is no longer being developed. EndNote, however, is an excellent package, available for Mac and PC, and they have a student pricing scheme which makes it quite cheap (about 80 Euro). Both packages will search PubMed for you too.

Finding your journal

You should aim your paper at a specific journal, even as you are starting to write up. Check your references: what journals carry this sort of article, or seem to have a special interest in the area? What other journals are there? There are many, many journals, so don't go downmarket until you have exhausted the better ones. There is a link below (under submitting a paper) to a clearing site for medical journals. It's useful for identifying journals you hadn't suspected of existing. Download the instructions for authors and follow them right from the start. It's easier than revising later.

Submitting a paper

Before you submit a paper, check carefully to see that it conforms to the journal's instructions for authors. If you need instructions for authors, the best single information resource is
http://www.mco.edu:80/lib/instr/libinsta.html
which lists over 2,000 journals. EndNote and ProCite comes with the formats for many journals built-in, which is another reason for having a proper bibliographic package.

If you use a bibliographic package such as Reference Manager, ProCite or EndNote to store your references, then the process of adding references in the journal's preferred format is straightforward. If you are adding references to your manuscript using footnotes within your word processor, changing from numbered references to author and date format can be very time consuming, so it is well worth identifying your target journal before you start referencing. If you write a lot of papers, a proper bibliographic database is a must.

Make sure that you have submitted enough copies of your manuscript, and that the letter of submission conforms to the journal's requirements (many journals require the letter to contain a copyright statement or a statement of participation from all the authors).

You will also need to write a covering letter. This should be short and direct. You should say briefly what your paper is about and why it is interesting. You should also try to make the case for publication in *this* journal rather than any other one. This is especially

important for general journals such as the BMJ and Lancet. They need to be convinced that your paper is of general interest, rather than being more suited to a specialist journal. If the journal has published papers in this area before, be sure to mention this. It establishes your paper as legitimate on the grounds of historical precedent, if nothing else!

I favour including supporting material when submitting. If you used a questionnaire, for instance, include copies so that the reviewers will have a clearer idea how you gathered your data. I also tend to include graphics which make the data easier to understand, with a note to the editor that they can be omitted if the referees think that the data would be better summarised numerically (graphics take up space, and editors tend to be cautious about having too many in an article). In fact, every time I have done this, the graphics have been published, but I have been prepared to back down.

Dealing with referees' comments

Don't panic!

Read the editor's letter first. This will often tell you the most important changes and clarifications that the referees are seeking. It may also contain hints as to which suggestions to ignore!

Then read and, if necessary, itemise all reviewers' comments. You will need to prepare a list of replies, based on this, detailing the changes you have made. It helps the editor and reviewers greatly if they can see your responses, item by item.

Always try to comply with suggestions to rephrase or rewrite sections that referees found hard to understand or misleading. This is the first time your paper has been read by a reader who didn't know the work as well as you did. Besides, it puts the referees in a good humour.

If you cannot comply with a suggestion (say, because you simply don't have the data), then put a bit into the paper mentioning the problem. For instance, if a reviewer has asked you to provide data from patient records, and you cannot, then you could put a sentence into the discussion saying *It is important to note that length of time between referral and appointment was based on the patients' recall. We could not verify the accuracy of this information because hospital confidentiality policy prevented us from accessing the relevant patient records.*

If you think a suggestion is off the wall, the editor probably will think so too. In that case, trust the editor, and put in a something like *Comment 8: We attempted to present the data in this way, but the resulting table was complex and confusing. We have therefore retained the original table, but with improvements in the captioning and layout which we hope will achieve the same effect as the referees' suggestion.* And make some cosmetic change to the paper to make the referee feel that they really did have a valid point. Never snub a referee directly.

If a referee is boorish and aggressive, you are actually at an advantage. By expressing

themselves rudely, referees embarrass journal editors (believe me, they really do!) and so the editor's sympathies will lie with you. All you have to do is be sweetly reasonable and the editor will love you. Remember to quote boorish reviewers verbatim — especially the really nasty bits. It will make the editor wince, and draw attention to the rudeness of the reviewer. But never, ever, be rude back. The ultimate pleasure is in seeing your work published, and hearing a gnashing of teeth out there.

And, above all, don't be put off by rejection. Good journals will often provide you with referees' comments even when they reject the paper. Read these carefully, but don't rush to revise. Make corrections and revisions that seem sensible to you, but only start on a major revision if there is a promise of publication. Most important: never leave a rejected paper lying about. Send it out to the next journal on the same day if you can.